Multisensory Emotional Recognition and Integration in the Ultra High Risk State and Early Phase of Psychosis
an fMRI Study

Tseng, Huai-Hsuan

Awarding institution:
King's College London

Multisensory Emotional Recognition and Integration in the Ultra High Risk State and Early Phase of Psychosis: an fMRI Study

Huai-Hsuan Tseng

MD, MSc

King's College London, Institute of Psychiatry

Thesis submitted to King's College London in partial fulfilment for the degree of Doctor of Philosophy (PhD)

September 2013

ABSTRACT

Patients in the early phase of psychosis show impairments of emotional processing. These patients also demonstrate neuroanatomical and neurofunctional abnormalities which are similar to those in patients with schizophrenia in regions that are involved in emotional processing. Impaired emotional processing is reported, albeit in an attenuated form, in individuals with an Ultra High Risk (UHR) for psychosis. To date however, few studies have specifically examined the neural substrate of emotional processing in the early and prodromal phase of psychosis.

The effective integration of emotional information is extremely important for interpersonal interactions and daily social functioning; but the disturbances of integration of multisensory emotional information and the associated neural processes in patients with the early phase of psychosis remain unclear. Moreover, there are no studies that have examined the integration of emotional information in the early and prodromal phase of psychosis.

To do this I developed a Multisensory Emotion Recognition and Integration Task (MERIT). In an fMRI experiment and examined the neural substrate for emotion recognition and multisensory integration and the possible alteration in sixteen UHR subjects and eighteen patients with first episode of psychosis (FEP), in contrast with twenty-one healthy controls (HC). FEP patients demonstrated impairments in both unisensory and multisensory emotion recognition, and reduced activation in the brain areas associated with emotional recognition. In UHR subjects, such alterations were less pronounced than in FEP patients. Both FEP and UHR groups did not show a significant alteration in the brain areas associated with integration, but FEP patients failed to

deactivate areas that may have been associated with irrelevant visual stimuli, and areas associated with the default mode brain network. A speculative model proposes that the posterior superior temporal area is important for integrating emotional information, and its activation can be modulated by modality-specific attention.

These results are in part consistent with the notion that, relative to HCs, FEP patients show neurofunctional alterations in emotional processing regions that are qualitatively similar to those previously observed in schizophrenia patients. UHR subjects showed altered behavioural performance and brain activation at an intermediate level between those in HC and FEP groups. This raises the possibility of establishing neurofunctional biomarkers for emotional processing that could be used to identify UHR subjects who have a higher risk of frank psychosis, a prospect which could be investigated in future prospective and longitudinal studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This work was funded by a Government Scholarship to study abroad by the Ministry of Education, Taiwan; and a King's International Graduate Scholarship (KINGS) award from King's College London, with stipend and study costs supported by a project grant from the NARSAD Grants for Independent Investigators to my principle supervisor, Dr Paul Allen.

I would like to acknowledge all the people who have provided supervision and guidance at all stages of the work towards my PhD, most notably, my principle supervisor, Dr Paul Allen, for his considered and insightful thoughts, constant all-round support, and generous allocation of time; my second supervisor, Professor Philip McGuire, for his consistently inspiring foresight and constructive advice; and also my subsidiary supervisor, Dr Jonathan Roiser in the Institute of Cognitive Neuroscience, for his comprehensive and accurate suggestions on the study design and analyses. It has been a privilege to work with all three of them. I would like to thank my former supervisor and senior colleague, Professor Shi-Kai Liu, for his academic suggestions and personal support through the PhD course.

I would also like to thank friends, colleagues and fellow PhD students whose encouragement, assistance and support throughout, have been invaluable. In particular, I would like to thank Dr Matthijs Bossong, Dr Gemma Modinos, Dr Irina Falkenberg, Dr Matthew Kempton, Dr William Pettersson-Yeo, Mr John G Mills, and Carly Samson.

I am exceptionally grateful to all those who gave up their time to participate the project as subjects. I would also like to thank those staff of the South London and Maudsley National Health Service Trust, in particular the Outreach and Support Intervention Services in South London (OASIS) for their invaluable help and collaboration with the recruitment process that made this study possible.

Lastly, I would like to thank my families. My wife Chia-Fen Hsu, for her love, tolerance and continuous support; my aunt, Tsui-Lein Tseng (曾翠蓮), my mother, Shue-Lan Lee (李雪蘭), their years of support and encouragement are beyond measure.

This thesis is dedicated to my grandmother, Tseng Chang Jin Tsun (曾張金蔥).

My Role in the Work Described

I am very grateful to my supervisors Dr Paul Allen, Professor Philip McGuire and Dr Jonathan Roiser who helped me to devise the original study design. For the MERIT task, I am grateful that Professor Steven Nowicki generously permit the use of DANVA2 in the UK, and I am in debt to Dr Jonathan Roiser and Dr Sunjeev Kamboj in UCL, and Mr Jeffery Dalton of CNS for designing the task. I am also grateful to my colleagues including Dr Irina Falkenberg, Dr Gemma Modinos and Ms Carly Samson, with whom I collaborated in order to recruit, assess and scan all participants included in the study. Dr Paul Allen, Dr Jonathan Roiser and Dr Matthijs Bossong taught and helped me to pre-process the neuroimaging data in preparation for subsequent analysis. Alongside these procedures, I personally i) conducted the systematic review (with helps from supervisors and Dr. Kuan-Ming Chen from National Normal University, Taiwan), ii) recruited and assessed all the healthy participants in the pilot and behavioural studies, iii) performed all preprocessing and analyses described in the thesis, iii) interpreted and discussed all results detailed in the thesis. Clinical assessments were performed by a number of trained psychiatrists allied either to the Outreach and Support in South East London (OASIS) service with respect to At-Risk Mental State participants, or, with respect first episode psychosis participants, any one of the early intervention in psychosis services offered by the South London and Maudsley (SLaM) National Health Services Trust located in Southwark, Lambeth, Lewisham or Croydon.

ABBREVIATIONS

AN(C)OVA	Analyses Of (Co-)Variance
BOLD	Blood Oxygen Level Dependent
CAARMS	Comprehensive Assessment of At-Risk Mental States
DANVA2-AP	Diagnostic Analysis of Nonverbal Accuracy 2 – Adult Prosody
DANVA2-AAP	Diagnostic Analysis of Nonverbal Accuracy 2 – African American Prosody
DASS-42	Depression and Anxiety Symptom Scale
DEER-T	Dynamic Emotional Expression Recognition Task
FEP	First Episode Psychosis
fMRI	Functional MRI
FWE	Family Wise Error
HC	Healthy Controls
LSAS	Liebowitz Social Anxiety Scale
MERIT	Multisensory emotional recognition and integration task
MNI	Montreal Neurologic Institute
MRI	Magnetic Resonance Imaging
PACE	Personal Assessment and Crisis Evaluation
PANSS	Positive and Negative Syndrome Scale
rAcc	Relative Accuracy
RM-ANOVA	Repeated Measures Analyses Of Variance
rRT	Relative Reaction Times
RT	Reaction Times
SPQ	Schizotypal Personality Questionnaire
SVC	Small Volume Correction
UHR	Ultra High Risk for psychosis
WAIS III	Wechsler Adult Intelligent Scale III
STG/STS/STC	Superior Temporal Gyrus/Sulcus/Cortex
MSI	Multisensory Integration
IFG	Inferior Frontal Gyrus
ERP	Event related potentials

Chapter 1

General Introduction

Psychosis is a devastating illness that has a huge impact on both interpersonal and social function. Emotional and social impairments are present before the onset of psychosis (Phillips and Seidman, 2008) and are associated with its later onset in people who are already vulnerable to the disorder (Cannon et al., 2008, Velthorst et al., 2011). Over the past two decades, an increasing volume of research has been conducted examining social and cognitive impairments

The early phase of psychosis comprises a high risk phase which may then be followed by a first episode of psychosis. Several terms have been applied to this high risk phase, including “At Risk Mental State” (ARMS), “Clinical High Risk” (CHR) or “Ultra High Risk” (UHR). The risk is associated with ‘attenuated’ psychotic symptoms (Yung et al., 2003a). At risk individuals may also have subtle subjective cognitive deficits, as well as affective disturbances, interpersonal difficulties and reduced occupational and educational function (Yung and McGorry, 1996). Though less severe, these clinical features are qualitatively similar to those seen in patients diagnosed with schizophrenia or other psychotic disorders (Jones et al., 1993, Yung and McGorry, 1996).

Unlike attenuated psychotic symptoms, which have been widely researched due to their necessity for an ultra high risk classification, emotional instability and prominent affective symptoms have been under-researched in the prodromal stage of psychosis

(Yung et al., 2003a). However, affective symptoms are Frequently observed in high risk subjects (Fusar-Poli et al., 2013a), and include disturbed emotional perception, altered emotional expression, and heightened emotionality (Phillips and Seidman, 2008).

Emotional perception and identification are crucial processes during social and emotional interactions. If there are problems at this stage the ensuing emotional experience and subsequent response are likely to be effected. In patients with schizophrenia, deficits of emotional recognition have been widely documented in both the visual and auditory sensory modalities (Edwards et al., 2002). Difficulties in emotion recognition have also been observed in at risk populations (Amminger et al., 2012a, Amminger et al., 2012b, Thompson et al., 2012) and patients in their first episode of psychosis (Pinkham et al., 2005, Reske et al., 2009, Brown and Cohen, 2010). However, despite social information in real life usually being delivered via multiple sensory modalities, the majority of studies on social cognition in schizophrenia have focused on facial emotion perception and identification only. To reflect the psychological function required in real social situations, emotion recognition ability across different sensory modalities needs further investigation.

Similar to unisensory emotion recognition, integration of emotional information from different sensory modalities is a complex but critical process that is necessary before reaching a social cognitive judgment. Individuals automatically process information from multiple sensory channels, and integrate the information into meaningful ideas and percept. Effective integration comprises extraction of important and coherent

information from different sensory modalities, and filtering out irrelevant information to form a holistic percept. This ability to integrate subtle affective messages relevant to social situations or cues, and recognise and make accurate inferences is crucial for understanding complex or ambiguous social situations. A deficit in this domain may underlie the interpersonal and social functional impairment observed in those who are at high risk of psychosis.

To provide a general overview of these topics in this chapter, I will start with an operational definition of FEP and the UHR state as used in this thesis. Next, I will review studies of emotional processing and its underlying neural correlates in schizophrenia, FEP and UHR subjects, with an emphasis on studies of emotion recognition. In addition, a brief introduction to studies of emotional integration is also provided (for a detailed systemic review, see chapter 2). Finally, I will outline the framework and objectives of this thesis, in the context of the background literature presented below.

1.1 Operational Definitions for the Research Cohort

1.1.1 Ultra-High Risk (UHR) for psychosis

The UHR is a relatively new concept. UHR subjects have an elevated risk of developing a psychotic disorder, relative to the general population. Estimates of the risk of transition to psychosis (1-3 years post diagnosis) range from between 16-40% (Yung et

al., 2003b, Ruhrmann et al., 2005, Yung et al., 2008b, Fusar-Poli et al., 2012). To identify UHR subjects, the current study used an assessment tool developed at the Personal Assessment and Crisis Evaluation (PACE) clinic in Melbourne, Australia(Yung et al., 1998) known as the he Comprehensive Assessment of At-Risk Mental States (CAARMS) (Phillips et al., 2000). To be considered as having an UHR for psychosis subjects must meet one of more of the following criteria: 1) Attenuated psychotic symptoms not of psychotic intensity (e.g. ideas of reference, odd beliefs or magical thinking, perceptual disturbance and/or paranoid ideation) occurring several times a week for between one week and 5 years, 2) Brief Limited Intermittent Psychotic symptoms (e.g. the same symptoms described for category 1, but of psychotic intensity, which have a duration of less than one week, and resolve spontaneously within that time) and/or, 3) Trait and state risk factors combined with a significant decline in cognitive and social functioning over the past year (e.g. the individual has schizotypal personality disorder, or a first degree relative with a psychiatric disorder or schizotypal personality disorder, combined with a significant decrease in mental state functioning that occurred within the past year, lasting between one month and 2 years and reflected by a 30 point reduction in the global assessment of functioning scale from premorbid level) (Yung, 2004).

For this thesis, the UHR subjects were recruited from Outreach And Support in South London (Broome et al., 2005, Fusar-Poli et al., 2013b) South London and Maudsley NHS Foundation Trust) between January 2012 and May 2013. The UHR diagnosis was validated by administering the CAARMS instrument. Cases with a diagnosis of organic psychosis were excluded (A detailed inclusion and exclusion criteria see Chapter 5).

1.1.2 First Episode Psychosis (FEP)

Studying patients with well-established schizophrenia in order to elucidate primary mechanisms of psychosis is not ideal, as a number of confounding factors, including exposure to anti-psychotic medication, and effects of illness chronicity may influence the findings. Investigating FEP patients allows researchers to reduce the effect of these potentially confounding factors.

In this thesis, FEP was operationally defined as 'first treatment contact'plus an ICD–10 diagnosis of psychosis (codes F20–F29 and F30–F33). Patients aged 18–35 years who presented with a first episode of psychosis to the South London & Maudsley Mental Health National Health Service (SLaM NHS) Foundation Trust between January 2012 and May 2013 were invited to participate the project shortly after their first episode of psychosis, with a clinical state of partial remission. The clinical diagnosis was validated by administering the Schedules for Clinical Assessment in Neuropsychiatry (WHO, 1992). Cases with a diagnosis of organic psychosis were excluded (a detailed inclusion and exclusion criteria see Chapter 5).

1.2 Emotional processing in UHR and FEP subjects

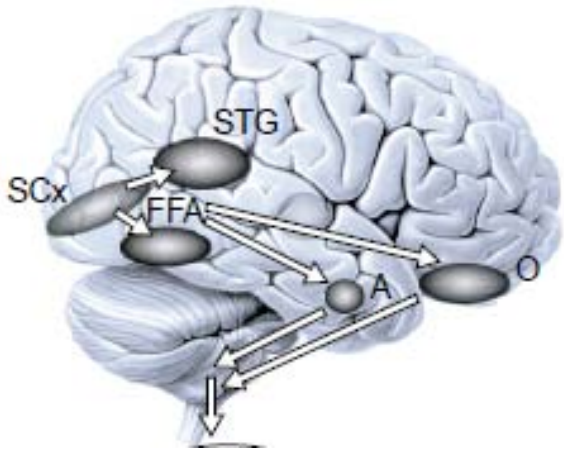
The historical definition of emotional processing describes "a process whereby emotional disturbances are absorbed, and that other experiences and behaviour can proceed without disruption" (Rachman, 1980). This definition encompasses several

stages of emotional processing: a state of emotional disturbance, the decline of the disturbance, and a return to normal, undisrupted behaviour. Later works focused on the cognitive process underlying the processing of emotional information and defined emotional processing as the modification of memory structures that underlie the information structure of emotion (Foa and Kozak, 1986).

Within the framework of studying emotional processing in schizophrenia however, the concept is adapted to describe domains which exhibit abnormalities in patients with schizophrenia, including impaired emotion perception, reduced self-reported positive experiences, reduced emotion expression and abnormal cognitive biases in the context of emotional information. Emotional processing difficulties are thus generally organised as difficulties in three domains: emotion perception, emotion experience, emotion expression (Kohler and Martin, 2006). A forth domain of emotional regulation has been suggested by Philips and Seidman (Phillips and Seidman, 2008) which may further describe the emotional instability which may partially be explained by the presence of cognitive biases. The current thesis is focused on the perception/recognition and integration aspect of emotional processing.

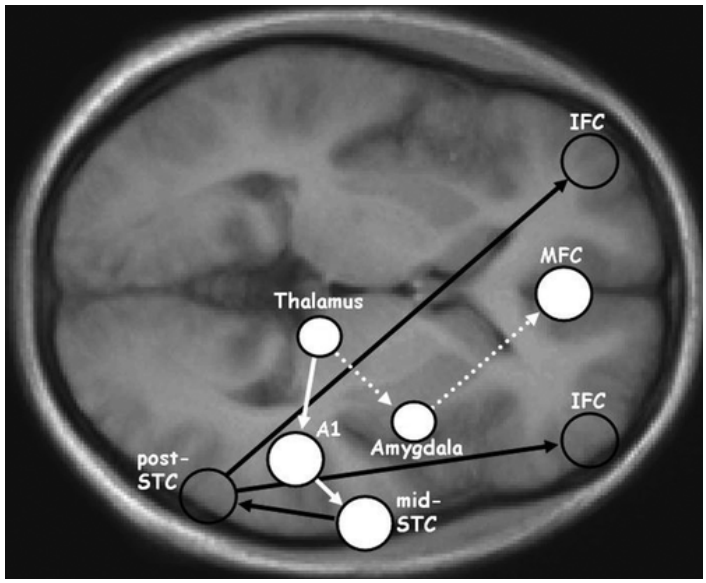
According to the model proposed by Adolphs (2002) and Wildgruber et al. (2009), the cognitive processing of emotional stimuli can roughly be divided into three stages: extracting and decoding modality-specific perceptual information; coarse labeling of the stimuli as expressing emotion, based on the decoded information; and fine

categorization according to conceptual knowledge (Aldophs, 2002, see Figure 1.1 and Figure 1.2).



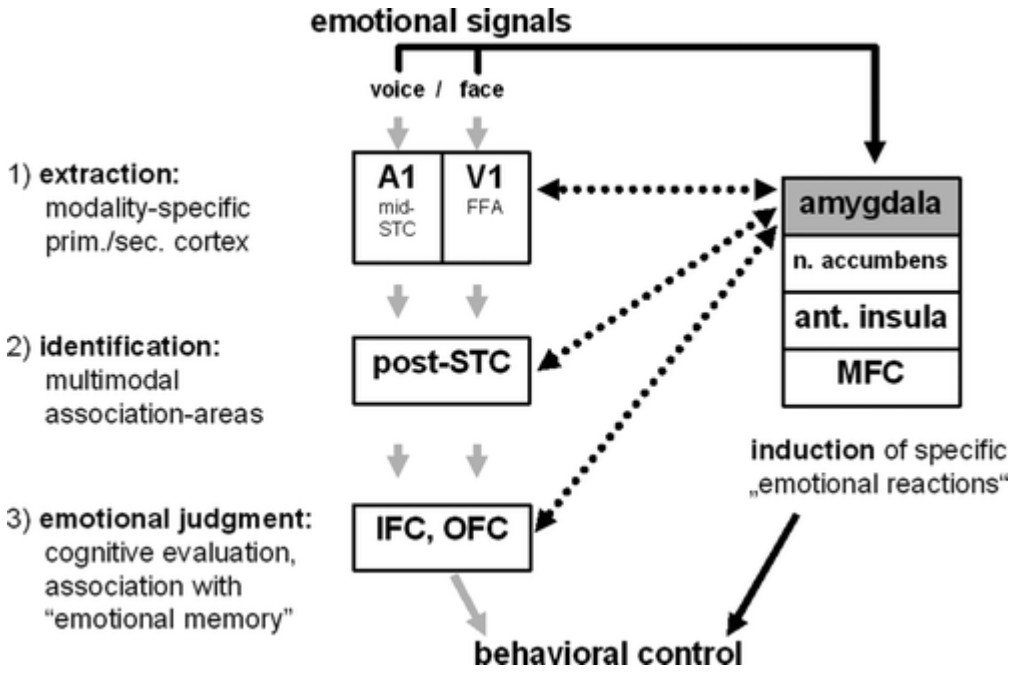
*STG = Superior Temporal Gyrus; FFA= Fusiform Area; A = Amygdala; O = Orbital frontal cortex; SCx = Striate cortex

Figure 1.1 Emotional face processing (Adapted from Aldophs, 2002)



*STC = Superior Temporal Cortex; A1= Primary Auditory Cortex; IFC = Inferior Frontal Cortex; MFC = Medial Frontal Cortex

Figure 1.2 Emotional prosody processing (Adapted from Wildgruber et al., 2009)



*STC = Superior Temporal Cortex; A1= Primary Auditory Cortex; V1 = Primary Visual Cortex; IFC = Inferior Frontal Cortex; MFC = Medial Frontal Cortex

Figure 1.3 An integrated model for emotional processing (Adapted from Wildgruber et al., 2009)

The first stage starts in the modality-specific primary/association cortex of the sensory percept: the fusiform gyrus in the visual modality, and the middle STG in the auditory modality (see Figure 1.3, (1)). During the second stage, although multiple brain areas are involved, posterior superior temporal sulcus/gyrus is considered important for the ‘emotional labeling’ for both visual and auditory information (see Figure 1.3, (2)). This area is also sensitive to human motions (Blake and Shiffrar, 2007), which is a crucial visual feature of emotion expressions.

Once the percept has been labeled as ‘emotional’, inferior frontal, orbital frontal areas and hippocampal structures will then be activated and trigger associated knowledge for the fine categorization (see Figure 1.3, (3)), including retrieval of conceptual knowledge

of emotion (Adolphs, 2002). At the same time, amygdala and limbic system may generate an emotional response in the subject (See Figure 1.3, 'Induction').

In this section, I will briefly review studies of emotional processing in UHR and FEP subjects to address the research question: *Are there emotional processing deficits, in particular, emotion recognition deficits, in UHR and FEP subjects?*

1.2.1 Emotional processing in schizophrenia and other psychotic disorders

Emotional dysfunction has been considered a hallmark of schizophrenia since its earliest conceptualization (Bleuler, 1911). In particular, an aberration in emotional processing is one of the core deficits in patients with schizophrenia (Meehl, 1962, Mandal et al., 1998). patients perform significantly worse on emotional recognition tasks than healthy volunteers, especially when negative emotional stimuli are used (Silver et al., 2002, Pinkham et al., 2003, Bediou et al., 2005), across a range of sensory modalities (Edwards et al., 2002). The ability to recognise and discriminate emotional stimuli is crucial for social cognitive processing and impairments in this domain may contribute to patients' poor interpersonal and social functioning (Hooker and Park, 2002, Pan et al., 2009).

The available literature suggests that there are positive associations between emotional processing difficulties and the severity of clinical symptoms in schizophrenia, particularly with negative symptoms (Schneider et al., 1995, Addington and Addington,

1998, Hoekert et al., 2007) (Chan et al., 2010). In addition, some evidence suggests that a state-dependant association (Kohler et al., 2000) may exist between the emotion recognition deficits and positive symptoms of schizophrenia (Shea et al., 2007) in both visual and auditory modalities (Kohler et al., 2000, Poole et al., 2000, Laroi et al., 2010). These deficits persist after adjusting for overall deficits in neurocognition, and also contributes to the formation and maintenance of delusions and hallucinations (Freeman and Garety, 2003).

1.2.2 Emotional processing in UHR and FEP subjects

Emotional dysfunction is also present in first-episode psychosis (FEP) patients and in individuals at clinical (Yung et al., 2003a, Addington and Addington, 2008), and familial high risk (Walker et al., 1993, Eack et al., 2010). Furthermore, factor analysis of high-risk symptoms yields a symptom cluster 'perceptual-affective instability' that includes marked emotional dysfunction, and high scores on this factor predict the subsequent onset of psychosis (Demjaha et al., 2010, Raballo et al., 2011). Hence, emotional dysfunction may be an early precursor to the development of positive symptoms and ultimately psychosis.

Deficits in facial and prosodic emotion recognition have been demonstrated in UHR populations (Addington et al., 2012, Amminger et al., 2012a, Amminger et al., 2012b, Thompson et al., 2012) and patients in their first episode of psychosis (FEP) (Pinkham et al., 2005, Reske et al., 2009, Brown and Cohen, 2010), with UHR subjects

demonstrating performance impairments intermediate between those in FEP patients and healthy controls (Thompson et al., 2012). This suggests that although the emotional recognition deficits in UHR population are slight, fundamental disturbances of emotional processing may already be apparent in those with attenuated symptoms.

1.3 The neural correlates of emotional processing FEP and UHR states

In this section, I will briefly review research investigating the neural correlates of emotional processing in FEP patients and UHR subjects to address the research question: *Are there alterations of neural correlates during emotional processing in FEP and UHR?*

1.3.1 The neural correlates of emotional processing in schizophrenia and other psychotic disorders

Schizophrenia is associated with neuroanatomical and neurofunctional abnormalities in emotional circuits, including the amygdala, hippocampus and the anterior cingulate cortex (Gur et al., 2002a, Hempel et al., 2003, Williams et al., 2004). These alterations are closely related to disturbances in cognitive and affective processing in schizophrenia (Habel et al., 2004). This thesis will focus on findings relating to the perception and recognition stage of emotional processing in different sensory modalities.

Patients with schizophrenia show significantly less activation in perceptual processing areas for both facial and prosodic emotional stimuli, including the right fusiform gyrus (Li et al., 2010) and left superior temporal gyrus (STG) (Mitchell et al., 2004). Activation in areas associated with emotional recognition and discrimination, including bilateral amygdala, parahippocampal gyrus and superior frontal gyrus, is also decreased when emotional face stimuli are used. Activation in these areas, as well as the patterns of lateralisation, also seem to be altered in schizophrenia patients during emotional prosodic recognition, although the direction and extent of these alterations is inconsistent across studies (Mitchell et al., 2004; Mitchell and Crow, 2005; Bach et al., 2009; Leitman et al., 2007; further details see chapter 6).

To summarise, it seems that the facial emotion processing difficulties are associated with an under recruitment of the amygdala and substantial hypoactivation throughout the ventral temporal-basal ganglia-prefrontal cortex, whilst prosodic emotion processing difficulties are associated with altered activation in temporal and parietal area and atypical lateralisation.

1.3.2 The neural correlates of emotional processing in FEP and UHR states

To date, the vast majority of research into neural alterations in UHR subjects has focused on cognitive rather that emotional deficits (Fusar-Poli et al., 2007b). Previous studies have also shown FEP patients and UHR subjects differ from healthy subjects

both in terms of their brain structure and function in many of the regions important for emotional processing (Fusar-Poli et al., 2007b, Smieskova et al., 2010).

So far only a few studies have explicitly focused on the neural substrate of emotional processing in the UHR subjects (Seiferth et al., 2008). Seiferth and colleagues showed hyperactivation in UHR subjects in the right lingual, fusiform and left middle occipital gyri during a facial emotion discrimination task. Another study of individuals who scored highly on a 'psychosis proneness' questionnaire showed altered prefronto-limbic functional connectivity during emotion regulation relative to subjects with low psychosis proneness (Modinos et al., 2010).

1.4 Multisensory emotional processing in Schizophrenia

In this section, I briefly summarised the studies of multisensory emotional processing in FEP patients and UHR subjects to address the research questions: *Is there a multisensory emotional processing deficit in FEP and UHR? If so, are there alterations of neural correlates during multisensory emotional processing in FEP and UHR?*

The prominent interpersonal and social impairments observed in patients with schizophrenia raises question about their ability to accurately perceive emotional and social information. Defective multisensory integration may be an underlying

mechanism related to emotional perception problems in schizophrenia. Multisensory integration (MSI) of non-emotional visual and auditory information is disturbed in schizophrenia (Surguladze et al., 2001, de Gelder et al., 2005, Ross et al., 2007). However, fewer studies have examined MSI of emotional information in schizophrenia. Although a diminished cross-modal influence of emotional faces on emotional voice categorization in patient with schizophrenia has been reported (de Jong et al., 2009), evidence in the field is inconclusive (de Gelder et al., 2003), and the possible mechanism related to MSI deficits remain speculative (de Jong, 2010; see Chapter 2).

Many functional imaging studies of multisensory emotional processing have been conducted in healthy subjects. During the presentation of congruent audiovisual emotional stimuli, areas responsible for audiovisual integration (superior temporal area), and emotional processing (inferior frontal gyrus, parahippocampal gyrus and the amygdala) are engaged (Park et al., 2010) (also see Chapter 2). During conflicting audiovisual emotional stimuli, a cingulate-fronto-parietal network is activated in healthy participants, likely associated with conflict monitoring and resolution (Muller et al., 2011).

To date, the neural substrate of multimodal emotional integration has not been fully examined in FEP patients or in UHR subjects. Areas responsible for higher-order audiovisual integration in healthy participants (Calvert et al., 2001, Ethofer et al., 2006b, Driver and Noesselt, 2008), located in the right STS and STG, are regions where functional and structural abnormalities are seen in patients with schizophrenia. In

addition, right inferior and middle frontal brain regions are also involved in audio visual integration and show altered engagement during audiovisual tasks in patients with schizophrenia (Ross et al., 2007).

1.5 Aims and Hypotheses

In summary, emotional processing difficulties are a robust feature of psychosis, and are evident in the first episode and in high risk subjects. The ability to integrate emotional information from different sensory modalities, consecutively, is an advanced social cognitive ability. Examining this complex social cognitive ability in FEP patients and UHR subjects may shed light on emotional processing impairments in the early phase of psychosis. However, relatively few studies have been conducted on emotional processing in FEP patients and UHR subjects, and no neuroimaging studies have been conducted examining the neural correlates on emotional integration in these groups.

The primary objectives and research questions addressed in the current thesis were as follows:

Research Question 1: Are there emotional processing and integration deficits in FEP patients and UHR subjects?

Hypothesis 1: Relative to healthy controls, FEP participants will demonstrate impaired recognition of single channel emotional information presented separately in the visual (facial expression; see section 5.1.4) and auditory (prosodic sentences; see section 6.1.4)

modalities. UHR subjects will demonstrate intermediate performance between controls and FEP.

Hypothesis 2: During a multisensory emotional recognition task, emotionally congruent multisensory stimuli (i.e. a fearful face presented together with a sentence of fearful prosody) will increase the accuracy of emotional recognition and reduce reaction times in healthy volunteers but not (or to a lesser extent) in FEP patients. In contrast, emotionally incongruent multisensory stimuli (i.e. a fearful face presented together with a sentence of happy prosody) will interfere with performance, subsequently decrease the accuracy of emotional recognition and increasing reaction times in healthy controls but not (or to a lesser extent) in FEP patients (See section 7.1.4). UHR subjects will demonstrate intermediate performance between controls and FEP in both facilitation and interference effects.

Research Question 2: Are there alterations in the neural correlates of emotional processing in FEP patients and UHR subjects?

Hypothesis 3: Relative to healthy controls, FEP participants will demonstrate altered regional brain activation during unimodality emotional information presented in visual (facial expression; see section 5.1.4) and auditory (prosodic sentences; see section 6.1.4) modalities separately, particularly in areas responsible for decoding facial and prosodic information, emotion recognition/identification and interpretation. UHR subjects will demonstrate alteration of activation intermediate between controls and FEP in both tasks.

Hypothesis 4: During a functional Magnetic Resonance Imaging (fMRI) study using the multisensory emotional recognition task, relative to healthy participants FEP participants will show altered activation in areas responsible for multisensory integration (superior temporal sulcus/gyrus), and emotion recognition/identification and interpretation, including amygdala, inferior frontal gyrus and hippocampal areas (See section 7.1.4). UHR subjects will demonstrate alteration of activation intermediate between controls and FEP patients.

1.6 Outlines of Chapters

In Chapter 2, I will systemically review both emotional and non-emotional studies of MSI in schizophrenia.

In Chapter 3, I will describe the process of designing and piloting a novel multisensory emotional recognition and integration task (MERIT).

In Chapter 4, I will report a study validating the MERIT in healthy participants (HC) and examine association between subjects' schizotypy scores and performance measures of the MERIT.

In Chapters 5 and 6, I will compare behavioural performance and brain activation patterns in HC, UHR subjects and FEP patients associated with the unisensory emotional processing of facial expression and prosodic voice separately to test Hypotheses 1 and 3.

In Chapter 7, I will examine task performance and brain activation pattern during the multisensory conditions of the MERIT in HC, UHR and subject groups to test Hypotheses 2 and 4.

Finally, a general discussion of the results and a speculative model explaining emotional processing and integration difficulties in psychosis will be presented in Chapter 8.

Chapter 2

A Systematic Review of Multisensory Integration in Schizophrenia

2.1 Introduction

Schizophrenia is a devastating disorder with onset usually in early adulthood and is associated with a range of social cognition and interpersonal deficits (e.g., Penn et al., 1997, Pinkham et al., 2003, Brittain et al., 2010). Impairments in these domains are closely related to the emotional processing deficits which are widely reported in patients with schizophrenia and other psychoses (Edwards et al., 2002). However, most social cognition studies in patients with schizophrenia have used emotional information presented in a single sensory modality, usually visual (Gur et al., 2002a, Kohler et al., 2003, Pinkham et al., 2005) or auditory (Leentjens et al., 1998, Leitman et al., 2005); for a comprehensive review, see Edwards et al. (2002). The etymology of schizophrenia means "splitting of mental functions", which implies poor functional integration of cognitive and affective aspects; however, less is known about how individuals with the illness integrate information from multiple sensory modalities.

We perceive real-time information from the environment through multiple sensory modalities (e.g. visual, auditory, and olfactory) automatically and simultaneously. At the same time information is filtered and composite relevant stimuli are integrated into meaningful ideas that initiate adaptive behaviours (Ethofer et al., 2006b). Furthermore,

the interplay between different sensory inputs generates a holistic experience, which differs from information available via any single sensory modality (McGurk and MacDonald, 1976). The underlying neurocognitive process that allows us to integrate streams of information from different sensory modalities into holistic experience is referred to as multisensory integration (MSI) or cross-modal influence.

MSI is a spontaneous, automatic process that occurs early when incoming perceptual information is adjacent temporally and/or spatially (de Jong et al., 2009), in both emotional and non-emotional contexts. A growth in MSI research in the last decade has explored integration in both temporal and spatial perceptual dimension. To this end, complex paradigms containing social, linguistics and emotional information have been developed to examine MSI. The information from different sensory modalities can be congruent, which typically facilitates processing and increases perceptual sensitivity and accuracy whilst shortening response latencies. This is usually referred to as a congruent facilitation effect (Miller, 1982, 1986, Schröger and Widmann, 1998). The information can also be conflicting along the dimension of judgment, which usually interferes with processing and decreases perceptual sensitivity and accuracy, and is referred to as an incongruent interference effect. When information is conflicting the judgment tends to be dominated by the most reliable source of input (Welch and Warren, 1980, Collignon et al., 2008), and sometimes leads to perceptual illusions, such as the McGurk effect (McGurk and MacDonald, 1976), in which conflicting visual speech influences auditory speech perception, or the ventriloquist effect (Choe et al., 1975, Bertelson and Radeau, 1981), in which misaligned visual cues disrupt auditory localization.

Compared to emotionally neutral stimuli, emotional stimuli have been shown to automatically capture attention. The integration of emotional information across modalities takes place before the semantic the and linguistic content of spoken words being processed (de Gelder and Vroomen, 2000). A congruent facilitation effect has been demonstrated while subjects perceive emotionally congruent information (de Gelder and Vroomen, 2000). In contrast, emotional interference or conflict effects are usually examined by presenting incongruent emotional cues in different sensory modalities (i.e., a sad facial expression with an angry voice) which results in decreased accuracy (de Gelder et al., 1999), and/or an increased reaction times (Dolan et al., 2001, Ochsner et al., 2009, Wittfoth et al., 2010).

Experimental findings in healthy volunteers suggest that multisensory integrative processes have an unavoidable nature (Muller et al., 2011), irrespective of the congruency or emotional content of the information. In schizophrenia, it is now recognized as an important area of research, as impaired MSI in patients is a clear disadvantage in a complicated, information-abundant world and is likely to impact on day-today social functioning. However, to what extent and in which domains this automatic integrative process is impaired in patients with schizophrenia is unclear. The aim of the current review is to provide a systematic overview of studies that examined multisensory integration in individuals with schizophrenia and schizophrenic spectrum disorders. We have included both non-emotional and emotional experimental studies.

2.2 Search strategy and results

An electronic search was performed using the PubMed database and the following search criteria: ("crossmodal" OR "audiovisual" OR "audio-visual" OR "multisensory" OR "multi-sensory") AND ("integration" OR "interference" OR "conflict") AND ("schizophrenia" OR "psychosis"). All studies published in the English language published or published on line before August 2013 were included. In addition, we manually searched the reference lists of the included studies to ensure that no studies of significance were omitted from the review. The literature search was repeated and the selection of the literature was cross-checked by one of the co-authors (Dr KM Chen). When a discrepancy in selection occurred, the thesis supervisors and Dr Chen were consulted to reach a consensus.

A total of forty-eight studies were identified. Nineteen studies that did not originally investigate multisensory integration in schizophrenic spectrum disorders were excluded (two review articles, nine articles focused on healthy populations, and eight articles that did not use multisensory integration as one of the main research paradigm). In patients with schizophrenia, fourteen studies investigated non-emotional MSI and seven studies examined emotional MSI (see Table 2.1). Furthermore, one study examining unusual proneness to visuo-tactile integration in schizophrenia, one multimodal case study investigating multisensory hallucinations, two post-mortem studies that assessed brain areas related to MSI, and four studies using a MSI task as part of a neurological examination battery in schizophrenia patients and their relatives were also included for

comprehensiveness (See Table 2.2). Overall twenty-nine studies were included in the current review (See Figure 2.1).

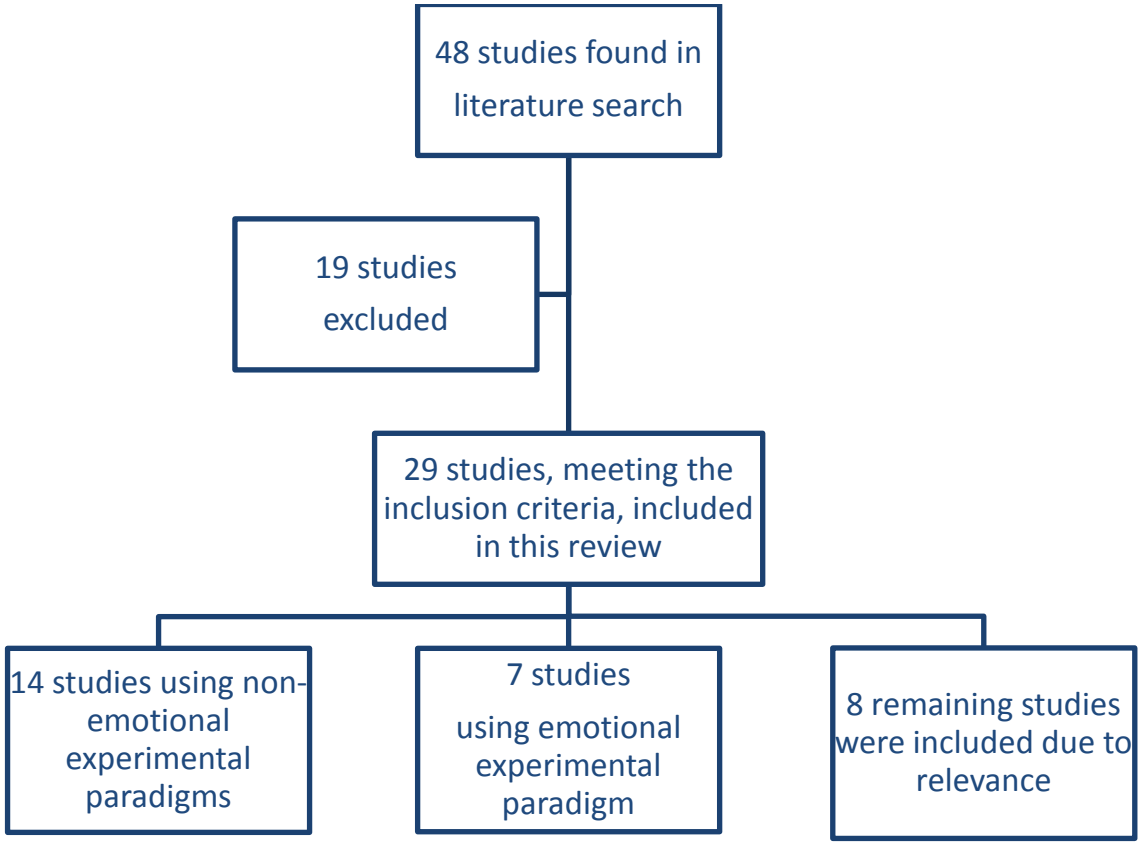


Figure 2.1 Flowchart of systematically reviewing protocol

2.2.1 Non-emotional multisensory integration in schizophrenia

Altogether, fourteen studies have been performed using non-emotional multisensory integration tasks comparing schizophrenia patients with healthy controls. Study

characteristics, sample sizes, task paradigm used and a summary of the results may be seen in Table 2.1(a).

Although four of the fourteen studies showed that patients with schizophrenia had a similar ability (Surguladze et al., 2001, de Gelder et al., 2003, de Boer-Schellekens et al., 2013), and two studies showed a compensatory increase in reaction time facilitation (Stone et al., 2011, Stephen et al., 2013) to process audiovisual stimuli while compared to healthy controls , the majority of studies demonstrated impaired performance in patients (de Gelder et al., 2003, Foucher et al., 2007, Ross et al., 2007, Pearl et al., 2009, Szycik et al., 2009, Williams et al., 2010, Stone et al., 2011, Szycik et al., 2013).

Generally, while using paradigms simultaneously presenting auditory and visual stimuli, healthy controls show shorter reaction times when processing congruent bimodal relative to unimodal information, demonstrating a congruent facilitation effect. This facilitation effect is reduced in schizophrenia patients (in terms of accuracy and RTs) during the presentation of congruent multisensory information (Williams et al., 2010).

Table 2.1 Summary of Studies of Multisensory Integration in Schizophrenia: (a) Non-emotional (b) Emotional

(a) Non-emotional

	Author	Method and Task paradigm	Sample size	
Non-emotional				
	de Gelder et al. (2003)	*Cognitive –Behavioural* *ventriloquist effect* *Lip-reading paradigm, incongruent*	18 SCH, 12 NC	No group differences in spatial multisensory integration ↓ lip-reading on audiovisual speech perception
	Pearl et al. (2009)	*Cognitive –Behavioural* *Lip-reading paradigm, incongruent*	30 SCH, 20 HC	↓ audiovisual integrative process in SCH
	Ross et al. (2007)	*Cognitive –Behavioural* *Lip-reading paradigm, congruent*	18 SCH, 18 HC	↓ benefit from congruent information in SCH
	Williams et al. (2010)	*Cognitive –Behavioural* *Target-detection paradigm*	20SCH, 20 HC	↓ benefit on RT for congruent information in SCH
	de Boer-Schellekens et al. (2013)	*Cognitive –Behavioural* *visual temporal order judgment task*	16 SCH, 16 HC	No deficits in the integration of low-level auditory and visual information ↓ visual temporal order judgement,
	Foucher et al. (2007)	*Cognitive –Behavioural* *simultaneity judgement paradigm*	30 SCH, 33 HC	No specific deficits in multisensory integration ↓ resolution of time perception
	Stone et al. (2011)	*ERP* *Far-near judgment*	14 SCH, 15 HC	↑ benefit from congruent information in schizophrenia ↑ absolute magnitude of evoked brain response from spatially congruent information in SCH
	Stekelenburg et al. (2013)	*EEG/ERP* *mixed lip-reading non lip-reading paradigm*	18 SCH, 18 HC	↓ suppression of auditory evoked brain response ↓ facilitation in reaction time in schizophrenia
	Stephen et al. (2013)	*Joint independent component analysis of MEG and DTI*	29 SCH, 29 HC	↑ RT facilitation benefit from congruent information in schizophrenia

	Far-near judgment		
Sass et al. (2013)	*fMRI* *semantic priming paradigm*	14 SCH, 14 HC,	↓ activation in superior frontal gyrus, middle temporal gyrus, anterior cingulate and angular gyrus SCH
Straube et al. (2013)	*fMRI* *Videos of an actor performing gestures in a concrete and abstract sentence context*	16 SCH, 16HC	↓ connectivity from STS to middle temporal gyrus and ventral IFG during the abstract relative to concrete gestures
Surguladze et al. (2001)	*fMRI and Behavioural* *Lip-reading paradigm, incongruent*	14 SCH, 7 HC	No group differences on audiovisual speech perception
Szycik et al. (2009)	*fMRI* *Lip-reading paradigm, congruent vs. incongruent*	15 SCH, 15 HC	Group × congruency: less activation difference or opposite activation pattern for incongruent than for congruent stimuli in SCH
Szycik et al. (2013)	*fMRI* *Lip-reading paradigm, congruent vs. incongruent*	15 SCH, 15HC	↓ adaptive connectivity in right posterior STS and left inferior frontal gyrus (Broca's area)

(b) Emotional

	Author	Method and Task paradigm	Sample size	
Emotional				
	de Gelder et al. (2005)	*Cognitive –Behavioural* *happy-sad face continuum* *face on voices* *voice on faces*	13 SCH, 13 HC	↓ influence of emotional voice on face continuum in SCH ↑ influence of emotional face on voice continuum in SCH
	de Jong et al. (2009)	*Cognitive –Behavioural* *happy-sad and happy-fearful face-voice pairs, congruent vs. incongruent* *face on voices*	50 SCH, 46 Non-SCH, HC	↓ benefit from congruent relative to incongruent information in SCH
	de Jong et al. (2010)	*Cognitive –Behavioural* *happy-sad and happy-fearful face-voice pairs with or without visual or auditory distractors, congruent vs. incongruent* *face on voices*	50 SCH, 46 Non-SCH psychotic	Comparing to conditions without distractors: ↓ benefit when with distractors in HC similar benefit when with visual distractors in SCH ↑ benefit when with auditory distractors in SCH
	Seubert et al. (2010)	*Cognitive –Behavioural* *pleasant and unpleasant olfactory priming on judging facial expression, congruent or incongruent*	24 SCH, 24 HC	↓ olfactory priming facilitation on disgust recognition in SCH, mainly unpleasant odour priming
	Van den Stock et al. (2011)	*Cognitive –Behavioural* *whole-body expressions with human or animal vocalizations, congruent or incongruent*	SCH, non-SCH psychotic	↑ cross-modality influence from human voices
	Muller et al. (2012)	*ERP* *happy, neutral, fear audiovisual emotional information, congruent or incongruent*	18 SCH, 18 HC	↓ amplitudes associated with visual processing during emotionally incongruent stimulus pairs
	Muller et al. (2013b)	*fMRI* *happy, neutral, fear audiovisual information, congruent or incongruent*	15 SCH, 15 HC	↓ deactivation of the left IPC during congruent audiovisual information

Table 2.2 Summary of other studies related to multisensory integration in schizophrenia

	Author	Method and Task paradigm	Sample size	
Others				
	Ferri et al. (2013)	*Rubber hand illusion paradigm*	21 SCH, 17 HC	↓Sense of ownership over the rubber hand
	Jardri et al. (2009)	*Case report, multimodal*	1 SCH	activation in bilateral STS, occipital temporal sulcus and anterior cingulate cortex during multisensory hallucination
	Smiley et al. (2012)	*Post-mortem study*	24 SCH, 24 Control	No differences in gray volumes of planum temporal; ↓ upper layer thickness of the caudal of planum temporal in SCH
	Smiley et al. (2009)	*Post-mortem study*	19 SCH, 18 Control	No differences in gray matter volume in inferior parietal lobule; ↓ thickness in angular gyrus relative to supramarginal gyrus in SCH
	Compton et al. (2006)	*Neurological examination includes audiovisual integration*	110 SCH	audiovisual integration task was included in a 'sensory integration factor'
	Sanders et al. (2000)	*Neurological examination includes audiovisual integration*	103SCH	audiovisual integration task was included in a 'cognitive-perceptual factor'
	Sanders et al. (2006)	*Neurological examination includes audiovisual integration*	96 relatives of SCH	audiovisual integration task showed statistically significant heritability
	Tumkaya et al. (2012)	*Neurological examination includes audiovisual integration*	30 SCH, 30 OCD, 16 SCH+OCD, 13 OCD with poor insight	performance of audiovisual integration task: SCH+OCD > OCD .

The reduced facilitation effect is further supported by findings from linguistic lip reading paradigms. Whilst healthy controls demonstrate an increase in accuracy when simultaneously processing congruent auditory and visual linguistic information, two studies show that patients with schizophrenia take less advantage of congruent audiovisual stimuli, again indicating a reduced facilitation effect (Ross et al., 2007, Szycik et al., 2009). Furthermore, whilst healthy controls demonstrate a decrease in accuracy for non-emotional conflicting information, accompanied with significant interaction responses of visual and auditory phonemes (McGurk effect), patients with schizophrenia show a less significant effect, indicating an attenuation of interference(de Gelder et al., 2003, Pearl et al., 2009).

Using a mixed linguistic (lip-reading) and non-linguistic (visual and auditory stimuli of hand clapping and tapping of a spoon) paradigm, a deficit of multisensory processing in schizophrenia is also seen at the neural level in evoked brain responses (Stekelenburg et al., 2013). When visual predictive information is presented, the auditory evoked brain response is suppressed and a decrease in reaction time is observed in healthy participants, but the suppression in amplitude is decreased and reaction times unchanged in patients with schizophrenia. The source location analyses showed deficits in the network subserving audiovisual integration, mainly right posterior STG, right middle temporal gyrus and left inferior frontal gyrus.

Contradicting to these findings, a study showed comparable ventriloquist effect (multisensory integration of spatially incongruent stimuli) for healthy participants and

patients with schizophrenia (de Gelder et al., 2003). Two studies examined multisensory interaction in schizophrenia using paradigms along the time domain. Both studies found that patient of schizophrenia were less able to discriminate consecutive stimuli as separate ones, but did not find a specific deficit in the integration of low-level non-linguistic multimodal information (Foucher et al., 2007, de Boer-Schellekens et al., 2013). By contrast, only one study using a linguistic lip-reading paradigm did not demonstrate MSI deficit in patients (Surguladze et al., 2001).

Furthermore, two studies show a compensatory increase in reaction time facilitation (i.e. the shortening of reaction time relative to single modality trials) in patient with schizophrenia, using a novel non-linguistic 'far-near' paradigm in both behavioural and evoked brain responses (Stone et al., 2011, Stephen et al., 2013). The authors suggested that this increase might be driven by greater unisensory deficits in patients with schizophrenia. By contrast, multisensory stimuli seem to capture their attention more readily than unisensory stimuli, consequently reducing their long RT associated with modality-specific attentional deficits during unisensory stimuli (Stephen et al., 2013).

The neurofunctional substrate for MSI in schizophrenia has been recently investigated in four functional Magnetic Resonance Imaging (fMRI) studies. During a lip-reading audiovisual task (using both congruent and incongruent stimuli pairs), greater activation was observed in multiple brain regions covering auditory, visual processing and attention functions for incongruent relative to congruent stimuli in healthy controls. A reversed activation pattern was seen in the right inferior frontal area and bilateral

superior and middle temporal gyrus in patient with schizophrenia, where greater activation for congruent relative to incongruent stimuli was observed. Cingulate, precuneus, parahippocampal and fusiform areas also showed similarly reversed activation pattern in schizophrenia (Szycik et al., 2009). During a semantic priming task, greater activation during multisensory versus unisensory stimuli was observed in superior frontal gyrus, middle temporal gyrus, anterior cingulate and angular gyrus in healthy controls relative to patients with schizophrenia (Sass et al., 2013). These results suggest that activation is altered in a network of sensory and language regions including the superior and middle temporal cortex, superior and inferior frontal gyrus in patients with schizophrenia across different multisensory paradigms (Szycik et al., 2009, Sass et al., 2013), with activation reduced relative to unisensory stimuli (Sass et al., 2013). The affected neural substrates may also involve cingulate cortex, angular gyrus and precuneus.

Further auxiliary findings provide anatomical evidence of the underlying neural substrate associated with MSI deficits in schizophrenia. In a case report, activation in bilateral STS, occipital temporal sulcus and anterior cingulate cortex was also reported during multisensory hallucination (Jardri et al., 2009); in concert with the post-mortem finding of thinner cortical layers of posterior temporal gyrus and angular gyrus in schizophrenia(Smiley et al., 2009, Smiley et al., 2012), support the notion that STS and the temporal-parietal junction are crucial for understanding the alteration of MSI in schizophrenia.

Connectivity between superior temporal sulcus (STS) and inferior frontal gyrus (IFG) was investigated in two fMRI studies using an audiovisual integration task. STS is functionally connected to disparate regions of the brain, but connectivity between STS and the middle temporal gyrus and ventral IFG is commonly reported (Straube et al., 2013). During a metaphoric gestures processing task, participants were asked to look at 'concrete' gestures that were mentioned in, or 'abstract' gestures that were associated with spoken words. Decreased connectivity from STS to middle temporal gyrus and ventral IFG in patients with schizophrenia during the abstract relative to concrete gestures was observed (Straube et al., 2013). Similarly, during a lip-reading audiovisual task, an altered connectivity pattern for congruent relative to incongruent trials was seen between the right posterior STS and left IFG (Broca's area) in patients relative to healthy controls. Less adaptive connectivity between left IFG and supplementary motor area were also observed (Szycik et al., 2013). These results suggest altered or aberrant connectivity between STS and IFG during MSI, particularly during incongruent stimuli and during integrating visual gestures in an abstract linguistic context in schizophrenia. This neurofunctional alteration might be the neural basis of interpersonal communication problems in patients with schizophrenia.

Overall, the body of evidence indicated that patients with schizophrenia have a deficit in integrating concurrent non-emotional audiovisual information, in terms of both reduced facilitation and interference effects. Studies using linguistic paradigms show an attenuated integration effect for patients with schizophrenia, although the results of studies using non-linguistic paradigms (de Gelder et al., 2003, Foucher et al., 2007, Williams et al., 2010, Stone et al., 2011, de Boer-Schellekens et al., 2013) are less clear.

Given these findings, de Gelder and colleagues (de Gelder et al., 2003) argue for a specific impairment in the audiovisual integration of cues with communicative value rather than a general impairment across perceptual domains. ERP and fMRI studies using linguistic paradigms support deficits of multisensory processing in schizophrenia at the neural level (Szycik et al., 2009, Sass et al., 2013, Stekelenburg et al., 2013, Straube et al., 2013, Szycik et al., 2013). The affected neural substrates mainly include posterior STG and STS, middle temporal gyrus and IFG, with altered activation patterns and connectivity particularly during incongruent and abstract conditions.

2.2.2 Emotional multisensory integration in schizophrenia

Seven studies have been published using emotional multisensory integration tasks in patients with schizophrenia. Study characteristics, sample sizes, the task used and a summary of the results are shown in Table 2.1(b) (Page 41).

Five of these studies used behavioural paradigms with complex emotional stimuli. Of these, four studies investigated audiovisual integration, three used facial and prosodic stimuli (de Gelder et al., 2005, de Jong et al., 2009, de Jong et al., 2010), and one used whole-body expressions with affective prosody (Van den Stock et al., 2011). The fifth study investigated olfactory-visual emotional integration (Seubert et al., 2010). The findings are inconsistent. Whilst there is evidence to suggest that patients with schizophrenia are less able to integrate emotional information across different sensory

modalities (de Gelder et al., 2005, de Jong et al., 2009, de Jong et al., 2010, Seubert et al., 2010), resulting in reduced facilitation effects for congruent conditions relative to unimodal or incongruent conditions, two out of five studies reported increased cross-modal influence in schizophrenia(de Gelder et al., 2005, Van den Stock et al., 2011). In particular, an increased facilitation effect has been reported using when facial expressions to improve categorization of emotional voices (de Gelder et al., 2005), as well as an increased cross-modal influence of vocal emotion (Van den Stock et al., 2011). In addition, in the presence of non-emotional auditory distractions, increased cross-modal influence of facial on vocal emotion information (de Jong et al., 2010) has also been observed in schizophrenia.

Two of these studies examined the neurofunctional basis of emotional MSI; one measured event related potentials (Muller et al., 2012), and the other is meta-analytic connectivity analysis (Muller et al., 2013b). Both studies applied the same experimental paradigm in which participants were asked to rate emotional and neutral faces whilst concurrently being distracted by emotional or neutral sounds. Compared with controls, patients with schizophrenia showed a significant decrease in amplitudes associated with visual processing during emotionally incongruent stimulus pairs, whereas similar amplitudes between groups were seen during congruent conditions. Similar deficits were also observed during a single-modality face condition (Muller et al., 2012). The meta-analytic connectivity modelling analyses (Muller et al., 2013b) examined the connectivity of inferior parietal cortex (IPC), including healthy controls, patients with depression and patients with schizophrenia. IPC is involved in memory, language and social cognition, and is deactivated during audiovisual congruent happy stimuli in

healthy controls (Muller et al., 2013a). Patients with schizophrenia failed to deactivate this area during congruent condition, which reflects the impairments in MSI, and are possibly associated with impaired memory and emotion processing in patients with schizophrenia. Both studies provide evidence of alteration or dysregulation of the neural substrates during emotional MSI in schizophrenia.

Overall, the behavioural and neurofunctional findings suggest altered integration of emotional information in patients with schizophrenia. Nevertheless, it is still unclear whether patients show reduced or increased bi-modal facilitation effects. It is possible that the sensory modalities of stimuli and the experimental context (for example, linguistic content or not, with or without unisensory performance as the contrast) have influenced these findings.

2.3 Discussion

This systematic review summarises studies that investigated multisensory integration in individuals with schizophrenia and schizophrenic spectrum disorders, both in an emotional and a non-emotional context. Additional studies which may shed light on the other aspects of multisensory interaction and neural substrates of MSI in schizophrenia were also included in this review.

A number of experimental paradigms have been used to examine multisensory integration in patients with schizophrenia, with various task complexities and perceptual domains. Overall, patients with schizophrenia showed deficits in integrating non-emotional audiovisual information, in terms of both reduced facilitation and interference effect. As most of the studies reported here have used linguistic paradigms, the results indicate that audiovisual integration of linguistic information in schizophrenia is impaired. However, there is less agreement between studies using non-linguistic paradigms (de Gelder et al., 2003, Foucher et al., 2007, Williams et al., 2010, Stone et al., 2011, de Boer-Schellekens et al., 2013).

Patients with schizophrenia seem to perform well when integrating simple audiovisual information (e.g. light, shapes, and sounds) along spatial and time domains (de Gelder et al., 2003, Foucher et al., 2007, de Boer-Schellekens et al., 2013). Though patients with schizophrenia are less able to discriminate consecutive stimuli across sensory modalities, this deficit appears to be secondary to a more generalised lower time-resolution problem (Foucher et al., 2007, de Boer-Schellekens et al., 2013). These findings indicate a domain-specific impairment in the integration of audiovisual cues with a communicative value rather than a general impairment across perceptual domains (de Gelder et al., 2003). Emotional information tends to be more complex and communicative in nature, and not surprisingly, patients with schizophrenia showed altered integration when processing emotional audiovisual information concurrently. However, there are apparent inconsistencies across studies with regard to the existence of an emotional MSI deficit in schizophrenia. More specifically, both decreases and

increases in the ability to integrate emotional information across different modalities have been reported.

The underlying mechanistic alterations associated with multisensory integration deficits in schizophrenia remain unclear. Higher attentional demands for multisensory tasks maybe required for patients with schizophrenia to perform MSI task (de Jong et al., 2010), thus pre-existing attentional deficits in patients with schizophrenia may decrease facilitation brought about by utilising additional information from a second sensory modality. Furthermore, modality-specific attention plays an important regulatory role while integrating multisensory emotional information and has been demonstrated to be deficient in patients with schizophrenia (de Jong et al., 2010). However, despite the existence of modality-specific attentional deficits, compensatory increase of facilitation effects during MSI are also reported, particularly in non-linguistic paradigms (Stone et al., 2011, Stephen et al., 2013). The interaction of attention-capturing nature of multisensory stimuli with pre-existing attention deficits in schizophrenia remains to debate.

Several possibilities may explain the discrepancy between emotional and non-emotional integration in schizophrenia. Unisensory emotional processing difficulties in schizophrenia have been robustly demonstrated in both visual or auditory modalities (See review of Edwards et al., 2002). Intuitively, basic emotional processing deficits in schizophrenia may alter the integration of emotional information. As such, the existence of emotional processing difficulties within single sensory modalities is likely to affect

the integration of emotional information across modalities. However, it is not yet clear whether the integration difficulties observed in schizophrenia are secondary or independent to impaired emotional perception and recognition in any single sensory modality. A possible way to address this issue would be to incorporate performance in single modality tasks as a covariate while examining the facilitation/interference effects in multisensory integration tasks, in order to exclude putative associations. Nevertheless, the possibility of alterations in emotional integration process per se and its underlying neural substrates still exist.

Another possible reason for the discrepancy is the heterogeneity between the experimental paradigms used. Most emotional MSI studies did not separately compare performance for congruent and incongruent multisensory stimuli to that of single modality stimuli, thus the reported cross-modal influence effect reflects the additive effect of congruent facilitation and incongruent interference. Although both effects reflect a degree of integration, incongruent stimuli are known to be more attention-demanding for conflict detection (Driver and Noesselt, 2008). In addition, incongruent stimuli activate a cingulate-fronto-parietal network involved in conflict monitoring and resolution (Muller et al., 2011), which is not usually observed in congruent multisensory integration. As a result, incongruent interference may be compounded by an attentional deficit in patients with schizophrenia.

An auditory-dominance effect during MSI in patient with schizophrenia may also explain inconsistent results. Patients might relay more on auditory information (de

Gelder et al., 2005, Ross et al., 2007, Van den Stock et al., 2011) while perceiving concurrent audiovisual information. Supporting evidence for this hypothesis includes an increased cross-modal influence of voices on emotional bodily expression (Van den Stock et al., 2011), but decreased influence of emotional faces on voices (de Jong et al., 2009) in schizophrenia patients. The near-normal integration of visual speech with voice speech reported by Surguladze and colleagues (Surguladze et al., 2001) also suggest dominance of voices in audiovisual integration. The auditory dominance hypothesis is, however, still under debate. The exaggerated cross-modal influence in schizophrenia is not only found in vocal on facial emotion, but also in facial on vocal emotion (de Gelder et al., 2005). Evidence also suggests that patients with schizophrenia rely less on the visual modality than healthy controls only for specific types of visual stimuli (Stone et al., 2011). A critical evaluation of the auditory predominance hypothesis is, thus, warranted. For example, comparing performance during incongruent stimuli to performance during unisensory visual and auditory stimuli separately could demonstrate a modality-specific interference effect.

The neural basis of multisensory integration deficits in schizophrenia has also been investigated. Converging evidence indicates that neurofunctional and anatomical alterations in the STS and superior temporal gyrus, areas responsible for higher-order speech integration(Calvert et al., 2001, Ethofer et al., 2006b, Driver and Noesselt, 2008), and the inferior frontal gyrus responsible for semantic and linguistic function (Szycik et al., 2009, Sass et al., 2013, Stekelenburg et al., 2013, Straube et al., 2013, Szycik et al., 2013) may be responsible for MSI deficits in schizophrenia. Areas in the default mode and attentional networks, including precuneus and cingulate gyrus, may also be affected

(Szycik et al., 2009, Sass et al., 2013, Szycik et al., 2013), leading to attentional problems in patients (de Jong et al., 2010). To date, the only neurotransmitter system which has been linked to multisensory integration is glutamate. Evidence for this comes from an animal model (Jacklin et al., 2012) which is beyond our scope of review.

Although an increasing body of research has sought to investigate possible deficits in multisensory integration in schizophrenia, most of the studies have focused on audiovisual integration. A number of potential confounding factors inherent to studies on patients with schizophrenia, such as illness chronicity and antipsychotic medications, may play a role in the observed results. In this regard, studies in patients with a first episode of psychosis, or in subjects with prodromal symptoms, may provide a unique opportunity to circumvent some of these limitations. Similar to patients with schizophrenia (Gur et al., 2002a, Hempel et al., 2003, Williams et al., 2004), patients with early psychosis and prodromal signs and symptoms show anatomical and functional abnormalities in the neural circuitry associated with emotional processing relative to healthy controls, such as in the right lingual and fusiform gyrus and prefronto-limbic functional connectivity (Seiferth et al., 2008, Modinos et al., 2010, Smieskova et al., 2010). Patients with schizophrenia and early/prodromal psychosis display functional alterations in these brain regions and this may affect their ability to effectively integrate multi-modal emotional information (Seiferth et al., 2008, Pauly et al., 2010). However, to date, no studies have examined the neural substrate of multisensory emotional processing in subjects with early or prodromal psychosis. Finally, the relationship between single and multimodality emotional deficits, and their neural underpinnings, remains unclear. Further studies investigating neural networks

involved in emotional processing of both single and multiple sensory modalities in patients with early or prodromal psychosis are thus warranted to clarify this issue.

2.4 Conclusion

Multisensory integration is a complex and important perceptual process. It is crucial in everyday life to integrate relevant information before generating prompt and appropriate behavioural responses in a complex and constantly changing environment. The available literature reviewed herein indicates that patients with schizophrenia demonstrate impairments in the integration of non-emotional audiovisual stimuli, especially for complex linguistic stimuli. Impaired MSI for emotional stimuli has also been reported in patients, although the direction of effect is unclear. MSI impairments may also exist in other sensory modalities, such as olfactory-visual integration. The underlying mechanisms of these deficits remain unclear, although an overarching attentional deficit in patients with schizophrenia may be an important factor. Such impairments are likely to be associated with the pronounced emotional processing and social cognition deficits in schizophrenia patients. Neuroimaging studies have shown that a network of regions including the posterior superior temporal cortex and inferior frontal gyrus play an important role during both multisensory non-emotional and emotional integration. Further neuroimaging research into MSI deficits in schizophrenia and early/prodromal psychosis is warranted to elucidate the neural mechanisms underlying this important deficit.

Chapter 3

Establish a Novel Tool for Multisensory Emotional Integration

3.1 Introduction

As detailed in Chapter 1 and Chapter 2, multisensory integration (MSI) refers to the process of integrating streams of information from different sensory modalities, adjacent temporally and/or spatially (Bertelson and de Gelder,2004; Calvert, 2004; de Gelder and Bertelson, 2003; de Jong, 2009). In order to generate adaptive behavioural responses, individuals need to filter out irrelevant or inconsistent information during the integrative process, so as to generate holistic experience relevant to social situations. Adequately perceiving and integrating subtle emotional information plays a crucial role in most of the social situations. As an important social cognitive function in everyday life, emotional MSI has been robustly examined in healthy participants. Though research in clinical cohort is scarce, an accumulating research interest has been focused on emotional MSI in patients with schizophrenia (e.g. de Gelder, 2005, de Jong, 2009, 2010, Muller, 2012).

An optimal task design demonstrating emotional MSI allows investigation for both the *facilitation effect*, whereby reaction times are decreased and recognition accuracy is improved; and the *interference effect*, whereby reaction times are increased and recognition accuracy is reduced. To enable the comparison of these effects, creating a

novel task comprising both congruent and incongruent multisensory emotional trials is mandatory. Moreover, further inclusion of the unisensory emotion recognition tasks is necessary for studies in clinical cohorts, for 1) as the baseline contrast for both congruent and incongruent emotional multisensory stimuli to show the facilitation and interference effect; 2) to measure the unisensory emotional recognition ability which may also be impaired in clinical samples and consequently influence the emotional MSI performance.

In this chapter, I briefly summarise the process of developing the novel task, *Multisensory Recognition and Integration Task (MERIT).* First of all, I summarise the two validated emotional tasks adapted for the current thesis, the *Dynamic Emotional Expression Recognition Task* for dynamic facial stimuli and *Diagnostic Analysis of Nonverbal Accuracy-2* for the prosodic voice stimuli. Secondly, I describe the pilot study for validating the prosodic trials in the UK population, followed by pilot study for validating the instruction to response. Finally, I report the pilot study to preliminarily examine the MSI effects using 15 pairs of congruent and incongruent multisensory trials. The validation of the complete task in healthy participants will be described in Chapter 4.

3.1.1 Facial Emotion Paradigm - Dynamic Emotional Expression Recognition Task

The facial expression paradigm was adapted from the Dynamic Emotional Expression Recognition Task (DEER-T; Platt *et al.*, 2010), and comprises forty-eight dynamically morphing emotional White-European faces that morph from a neutral to an emotional expression, or between two neutral expressions (mouth open to mouth closed). The

stimuli are from a validated series of facial affect pictures originally developed by Tottenham et al. (Tottenham et al., 2009).

Dynamic emotional stimuli were created with Abrosoft Fantamorph software (version 4.0). Photographs of neutral expressions were morphed with photographs of happy, sad, fear and neutral expressions, with 12 morphing video clips in each emotional category. Each 'morph' comprises 25 frames and has a total duration of 3000 milliseconds. Over the duration of the video clip the target emotion became increasingly intense (see Figure 3.1). The task shows good psychometric properties (internal consistency) with α values of 0.97-0.98 (Kamboj et al., 2012).

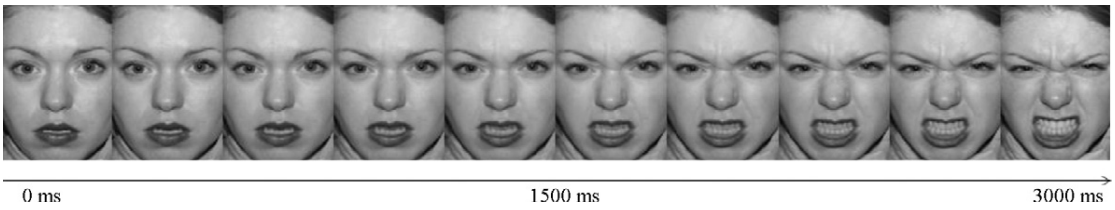


Figure 3.1 An example of emotional morphing faces

3.1.2 Voice Prosody Paradigm – Diagnostic Analysis of Nonverbal Accuracy – Adult Prosody (DANVA2-AP) and African American Prosody (DANVA2-AAP)

This paradigm comprises twenty-four adult prosody subtest trials adapted from the Diagnostic Analysis of Nonverbal Accuracy 2 – Adult Prosody subtest (DANVA2-AP) (Nowicki and Duke, 1994) and Diagnostic Analysis of Nonverbal Accuracy 2 – African American Prosody subtest (DANVA2-AAP) (Baum and Nowicki, 1998) with clips spoken by two male and two female professional American actors (one European and

one African within each gender). All voice clips use the following semantically neutral statement with happy, sad, and fearful prosodic tones: “I’m going out of the room now, but I’ll be back later”. The subtests consist of prosodic clips with 80% agreement for type and intensity of emotion rated by 146 adults.

3.2 Pilot study to validate American emotional prosody in UK English-speaking population

The DANVA2 has been validated in northern American English speaking populations (Nowicki, 2012) and recently, Australian English speaking populations (Thompson, 2012). However, due to the difference in accent and culture, a pilot study was conducted to validate DANVA2-AP and DANVA2-AAP in a small UK population. 21 healthy volunteers whose first language is English and whom were born in the British Isle were recruited from staff and students at King’s College London and University College of London (11 male, 10 female, mean age = 24.3 years old). They are asked to judge the emotional prosody of DANVA2-AP and DANVA2-AAP voice clips. Because reaction time is an important measure of the facilitation and interference effect of cross-modal integration, we tested two response options; 1) where participants were instructed to react as soon as they were confident they recognised the prosody DURING the voice clip and 2) where participants were instructed, to respond as soon as possible AFTER the voice clips finished. The former was likely to be a more accurate measure of the reaction time, and the later is the standard paradigm used in most prosodic emotion recognition paradigm, including DANVA2.

3.2.1 Recognition accuracy in a UK population

The overall accuracy of DANVA2-AP and DANVA2-AAP was 0.71 ± 0.08, which is significantly lower than the accuracy reported in American population (0.78 ± 0.09) sample (Nowicki, 1994). Based on the same selection criteria as DANVA2 (i.e. high agreement level but not extremely high so as to avoid the ceiling effect), the current study examined the agreement level of each prosodic voice clips in a British sample whose first language was English and who were born in the British Isles. Most of the items had similar high agreement rate, but some voice clips had lower agreement rates. Based on the pilot study results, 24 out of 48 clips with high agreement levels were selected, 8 for each emotional category (happy, sad and fear). These voice clips will be used in the fMRI paradigm. The difference of accuracy across the selected items was non-significant between UK and American sample (See Table 3.1).

	Pilot (N=21) Mean ± SD	**Original American Sample Mean ± SD (N=174)**	**t**
Overall	0.71 ± 0.08	0.78 ± 0.09	3.40***
Happy	0.69 ± 0.19	0.83 ± 0.15	3.92***
Sad	0.79 ± 0.12	0.84 ± 0.15	1.47
Angry	0.67 ± 0.14	0.76 ± 0.17	2.33*
Fearful	0.70 ± 0.20	0.68 ± 0.20	0.43
Selected items	0.76± 0.18	0.78 ± 0.09	0.84

* $p<0.05$; *** $p<0.001$

Table 3.1 The result of DANVA2-AP and DANVA2-AAP prosody recognition in native English-speaking participants in UK

3.2.2 Comparison between response instructions

The overall accuracy when participants were instructed to respond DURING the voice clips was 0.66 ± 0.15. Overall, this was not significantly different to when participants were instructed to respond AFTER the voice clips had finished. However, the accuracy of angry voice clips was significantly lower for the DURING compared to the AFTER instruction. There were no significant differences for other emotions (See Table 3.2). These results indicate that using the 'DURING' instruction are comparable with the classic 'AFTER' paradigm if 'angry' emotion voice clips are removed from the fMRI paradigm.

	Current Paradigm (DURING) Mean ± SD (N=21)	**DANVA2 Paradigm (AFTER) Mean ± SD (N=21)**	**t**
Voice Overall	0.66 ± 0.15	0.71 ± 0.08	1.35
Happy	0.69 ± 0.16	0.69 ± 0.19	0.00
Sad	0.73 ± 0.09	0.79 ± 0.12	1.83
Angry	0.55 ± 0.16	0.67 ± 0.14	2.59*
Fearful	0.65 ± 0.15	0.70 ± 0.20	0.92
Selected items	0.74± 0.15	0.76 ± 0.18	0.39

* $p<0.05$

Table 3.2 Comparison between Current (DURING) and Original DANVA2 Paradigm (AFTER)

3.3 Pilot study of multisensory emotional integration and interference in UK native English speakers

Ten volunteers (4 male, 6 females) were further invited for this part of pilot study. Based on the results of the previous pilot study, it was decided to use the 'DURING' response paradigm only in the multisensory task. In this part of the study, 15 pairs of morphing faces adapted from the DEER-T were displayed concurrently with 15 DANVA2-AP and DANVA2-AAP voice clips; both congruent and incongruent pair trials were included. Participants were asked to pay attention on both the morphing facial expression and voice clips, but to base their judgments on the prosody of the voices clips. Accuracy and reaction times were recorded as a measure of facilitation and interference effects.

3.3.1 Facilitation during emotionally congruent trials

The pilot results show that the mean accuracy for congruent trials showed a trend towards facilitation effect (0.83±0.31) in comparison with the average accuracy in unisensory prosodic trials (0.74± 0.15). The mean reaction time during congruent trials (2090±780 milliseconds) was similar to those for the prosodic unisensory trials (2230±310 milliseconds).

3.3.2 Interference during emotionally incongruent trials

The mean accuracy for incongruent trials (0.63±0.49) was numerically less than the mean accuracy for single modality voice clips (0.74± 0.15), but not statistically significant. The mean reaction time during congruent trials (2380±810 milliseconds) was similar to those for the prosodic unisensory trials (2230±310 milliseconds).

3.3.3 Congruent versus Incongruent trials

The mean accuracy for congruent trials was numerically greater, and the reaction times were numerically faster than those for the incongruent trials, though not statistically significant.

However, significant difference can be found while inspecting different emotional categories. During happy and fear trials, the accuracy for congruent pairs was significantly greater than for incongruent ones, and a trend of shorter reaction time during congruent sad trials comparing with incongruent sad trials (See Table 3.3).

These results indicate that there are at least trend of MSI effects can be observed within this small sample, with the expected direction of the effects, that is, more accurate and faster during the congruent trials, and less accurate and slower during the incongruent trials relative to the unisensory prosodic trials.

Emotional Category	**Accuracy** ***Congruent***	**Accuracy** ***Incongruent***	**Emotional Category**	**RT(ms)** ***Congruent***	**RT(ms)** ***Incongruent***
Happy[*]	1.00 ± 0.00	0.61 ± 0.49	**Happy**	2227 ± 488	2408 ± 620
Sad	0.56 ± 0.53	0.50 ± 0.52	**Sad**[+]	2050 ± 565	2605 ± 775
Angry	0.78 ± 0.44	0.78 ± 0.43	**Angry**	2182 ± 1148	2102 ± 1149
Fearful[*]	1.00 ± 0.00	0.63 ± 0.49	**Fearful**	1918 ± 860	2373 ± 786
Overall	0.83±0.31	0.63 ± 0.49	**Overall**	2090 ± 780	2380 ± 810

[+] $p<0.10$; [*] $p<0.05$

Table 3.3 Accuracy and Reaction Time in incongruent audiovisual emotional pairs under the instruction of based on voices

3.4 Conclusions

In sum, the current unisensory emotional paradigms show similar unisensory accuracy in UK population, after removing items with lower accuracy in UK population and the angry emotion. The current multisensory paradigm shows satisfactory potential to demonstrate MSI effects. The complete unisensory and multisensory trials of MERIT were then generated accordingly. A full description of the MERIT can be found in Section 4.3.2, page 70).

Chapter 4

Emotional Multisensory Integration in psychometric schizotypy

4.1 Introduction

Multisensory integration (MSI) refers to the process of integrating streams of information from different sensory modalities, adjacent temporally and/or spatially (de Jong et al., 2009). A stream of sensory of information from the environment enters awareness through multiple modalities (visual, auditory, tactile etc). Despite this, individuals normally perceive information from different sensory channels automatically, simultaneously and holistically. Specifically, individuals need to filter out irrelevant information, and integrate the relevant information into meaningful ideas in order to generate adaptive behavioural responses. The integration of information from different sensory inputs generates an unified experience that is not provided by any single modality (McGurk and MacDonald, 1976).This process occurs in both non-emotional and emotional contexts. A facilitation effect, whereby reaction times are decreased and recognition accuracy is improved, has been demonstrated when healthy individuals perceive emotionally congruent information from different sensory modalities (de Gelder et al., 1999). Conversely, when incongruent emotional information is presented (e.g., a sad face with an angry voice), interference effects are evident, with slower reaction times and reduced recognition accuracy (Dolan et al., 2001, Ochsner et al., 2009, Wittfoth et al., 2010).

Though MSI has not been studied extensively in clinical samples, there is evidence that patients with schizophrenia have difficulties integrating both emotional (de Gelder et al., 2005, de Jong et al., 2009, de Jong et al., 2010, Seubert et al., 2010) and non-emotional information (de Gelder et al., 2003, Ross et al., 2007, Pearl et al., 2009, Williams et al., 2010). For example, patients with schizophrenia have difficulty integrating information at both the perceptual and cognitive levels (de Gelder et al., 2003, Williams et al., 2010) and demonstrate attenuated cross-modal facilitation effects during the presentation of congruent emotional stimuli during emotional categorization (de Gelder et al., 2005, de Jong et al., 2009, de Jong et al., 2010). However, impaired emotional MSI in schizophrenia could reflect impairments in early unisensory processing (de Jong et al., 2010), and unimodal emotional processing deficits (Edwards et al., 2002, Kohler et al., 2010).

Schizotypy is a heritable, multidimensional construct related to schizophrenia, referring to a personality style that is associated with particular cognitive styles and unusual perceptual experiences (Claridge et al., 1985). The cognitive style includes unusual beliefs and peculiar ideas, as well as eccentric or odd behaviours, and social awkwardness or aversion [17-19]. Schizotypy is associated with an increased risk for developing schizophrenia (Kendler et al., 1993, Chapman et al., 1994).

Emotion processing difficulties are also evident in individuals with schizotypy (Brown and Cohen, 2010, Abbott and Green, 2012). Given that deficits in emotion information

processing skills are related to impairments in social functioning(Hooker and Park, 2002, Pan et al., 2009), and higher rates of interpersonal psychopathology (Abbott and Green, 2012), these deficits may reflect an important aspect of continuum spectrum between schizotypy and schizophrenia. A limitation of these studies however, is that they primarily investigated sensory modalities in isolation, whereas real social situations require individuals to process and integrate information from different sensory modalities. It is unknown whether schizotypy is associated with emotional integration difficulties. However, de Jong and colleagues reported that, compared with non-schizophrenia psychosis patients and healthy volunteers, patients with schizophrenia demonstrated a reduced facilitation effect during congruent emotional trials relative to incongruent trials (de Jong et al., 2009). These result suggest an impaired cross-modal integration of emotional information, which is in line with other studies in patient of schizophrenia using audiovisual (de Gelder et al., 2005, Van den Stock et al., 2011) or olfactory-visual paradigms (Seubert et al., 2010). However, the study did not examine accuracy for single modality emotional recognition, and as such, the results may be confounded by poor emotional recognition in these patients.

A bias towards negative emotional information may also influence multimodal sensory integration in schizotypal individuals. Like patients with schizophrenia, individuals with higher levels of schizotypy have been reported to experience more negative and less positive affect (Phillips and Seidman, 2008), and also exhibit negative biases in unimodal emotion recognition (Demenescu et al., 2010). Accordingly, during a cross-modal emotional integration task, individuals with higher schizotypal traits may negatively interpret positive emotional information, resulting in decreased recognition

accuracy regardless of the facilitation or interference effects of emotional stimuli presented in a second sensory modality. As a corollary of this effect, accuracy for negative emotional information should be relatively unaffected by interference effects from incongruent positive emotional information presented in a second sensory modality.

The aims of the present study were two-fold. First, we sought to examine emotional congruent facilitation and incongruent interference effects in healthy volunteers using a novel cross-modal emotional processing task, which employs dynamic facial affect and prosodic sentence stimuli. Second, we tested the hypotheses that schizotypal personality traits influence the ability to integrate emotional information, and bias individuals towards emotional information.

4.2 Predictions

First, we predicted that congruent emotional stimuli (e.g. a fearful face presented together with a spoken sentence of fearful prosody) would facilitate emotional recognition, as indexed by increased reponse accuracy and faster reaction times. Conversely, we predicted that incongruent emotional stimuli (e.g. a fearful face presented together with a spoken sentence of happy prosody) would interfere with emotional recognition, decreasing accuracy and increasing reaction times.

Second, we predicted that schizotypal traits would be associated with a reduced ability to integrate emotional information, resulting in a reduction in the facilitation and interference effects seen in non-schizotypal subjects.

Finally, we predicted that schizotypal traits would be associated with a negative information processing bias resulting in reduced accuracy and increased reaction times for positive emotional stimuli relative to negative stimuli during both single and multimodal emotional trials.

4.3 Methods

4.3.1 Participants

We studied twenty-five healthy participants (14 females, mean age 22.0, ranging from 18 to 34, S.D. 4.1; all right handed) without any history of psychiatric illness. The participants were screened using the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) for past or present psychopathology. Participants were required to have English as their first language as all the instructions and prosodic materials in the experiment were presented in English. People with a history of neurological illness or head injury requiring hospitalisation, or participants reporting excessive use of alcohol based on recommended safe limits suggested in the current guidance from the Royal College of Psychiatrists in the UK (>21 units per week for men and >14 units per week for women) or recent use (in the past two weeks) or regular use (more than once a week over 6 months in the past year) of recreational drugs (cannabis, stimulants,

hallucinogens, or opiates) were excluded. The study obtained ethical approval from the local Research Ethics Committee (reference number: 11/LO/0623) and written informed consent was collected from all participants.

4.3.2 Multisensory Emotional Recognition-Integration Task (MERIT)

First, participants completed a block of single modality trials, comprising the 48 stimuli from the DEER-T and the 24 stimuli from the DANVA2 AP and DANVA2-AAP, randomly interspersed. The inter-trial interval was 1 second, and the overall task duration was around six minutes. Then, during the MERIT, voice stimuli from the DANVA2-AP and DANVA2-AAP and morphing face video stimuli from the DEER-T were presented in pairs simultaneously. In each audiovisual pair, the length of the morphing process was adjusted to the exact length of the paired voice clips. The emotional face and voice clips were paired in a factorial design with three factors: congruency (congruent and incongruent multi-sensory trials), target modality (face and voice) and target emotion (happy, sad and fear). There were 48 pairs of stimuli in total, with 4 stimulus pairs in each of 12 conditions: 3 congruent conditions; 6 incongruent conditions; and 3 neutral (face)-emotional pairs. Data from the neutral-emotional pairs were not analysed further. On all trials, participants were asked to respond as quickly as they felt confident about the judgment before the end of the clip.

Two counterbalanced dual-modality blocks, the “Face” experiment and the “Voice” experiment were presented to participants. Participants were instructed to attend to the emotions within the faces and voices as a whole, but to base their response on the

emotional information in the faces (during the face experiment) or the emotional information in the voice (during voice experiment) when they felt that the emotions were different (i.e. during incongruent trials).

In both the Face and Voice experiments, stimuli were presented on a 17 inch laptop screen using a program developed by the research team under a Windows 7 32-bit environment. Participants were instructed to choose between four emotional categories (happy, sad, fear, and neutral) via a keyboard as quickly as possible before the voice and/or video clips ended, and accuracy and reaction times (RT in milliseconds) were recorded.

4.3.3 Outcome measures

We calculated the difference in accuracy between dual-modality congruent and incongruent trials relative to single modality trials. Thus, **relative accuracy (*rAcc*)** demonstrates either a facilitation effect (a positive value indicates facilitation: higher accuracy relative to single-modality trials) or an interference effect (a negative value indicates interference: lower accuracy relative to single-modality trials)**.** We expected an increase in accuracy for the congruent trials (facilitation effect) and a decrease in accuracy for the incongruent trials (interference effect). Similar calculations were performed for RTs to calculate **relative reaction time (*rRT*)** scores**,** demonstrating either a facilitation effect (a negative value indicates facilitation: shorter RT relative to single-modality trials) or an interference effect (a positive value indicates interference: longer RT relative to single-modality trials)**.** We expected shorter RTs for the congruent

trials (facilitation effect) and longer RTs for the incongruent trials (interference effect). Relative scores for accuracy and RT were averaged across the face and voice experiments. *rAcc* and *rRT* were also calculated separately for positive (happy) and negative (averaging sad and fear) target emotion trials.

4.3.4 Questionnaires

All participants completed the Schizotypal Personality Questionnaire (SPQ)(Raine, 1991) to evaluate levels of schizotypal personality traits and the Depression and Anxiety Symptom Scale (DASS-42) [9] to evaluate current affective states. The DASS-42 provides separate measures of anxiety, depression and stress along with a compound total score, which was used as an estimate for overall negative influence from negative affect.

4.4 Data Analysis

Outliers of reaction times, which were defined as 2.5 standard deviations above or below their mean reaction times for a particular condition, have been screened and checked for not significantly influence the average reaction times. Accuracy and RT were analysed using separate repeated measures analyses of variance (RM-ANOVAs). Separate RM-ANOVAs were also conducted for congruent and incongruent trials resulting in four models in total. Single/Multi-sensory modality (dual (i.e. congruent or incongruent) relative to single-modality trials), target modality (face experiment, voice experiment) and target emotional category (happy, sad, fear) were entered as within-

subject variables. Post-hoc paired *t* tests with Bonferroni correction for three emotions were used to clarify significant main effects and interactions involving emotion.

Pearson's correlation analyses were performed to examine the relationship between the level of schizotypy (SPQ) and the degree of integration effect, examining both accuracy (*rAcc*) and RTs (*rRT*). To investigate the influence of schizotypy on negatively biased information processing, we also calculated the difference scores of negative (sad, fear) stimuli relative to positive (happy) stimuli regarding both accuracy (dAcc [negative-positive]) and RT (dRT [negative-positive]) for both single and dual modality trials to indicate the negative biases toward incoming emotional information. Correlation analyses were then performed between SPQ scores and these difference scores.

4.5 Results

4.5.1 Psychometric measures

Demographic and psychometric data are shown in Table 4.1. A positive correlation was observed between SPQ and DASS scores (r=0.50, p<0.05), especially the depression subscale of the DASS (r=0.61, p<0.01). Accuracy and RTs of Baseline (Single Modality), congruent and incongruent trials are shown in Tables 4.2 and 4.3, respectively.

	Mean ± S.D. (Number)
Demographic variables	
Gender	M/F :11/14
Current age (y/o)	22.04 ± 4.10
Questionnaires	
SPQ	12.26 ± 8.58
DASS Total	9.92 ± 9.19
Depression	3.04 ± 4.46
Anxiety	1.80 ± 2.80
Stress	5.08 ± 4.44

Table 4.1 Demographic background and results of Questionnaire

	Face Accuracy Mean ± SD	Face RT Mean ± SD	Voice Accuracy Mean ± SD	Voice RT Mean ± SD (milliseconds)
Overall	0.87 ± 0.34	2145.33 ± 795.15	0.70 ± 0.46	2725 ± 1043
Happy	0.89 ± 0.31	1741.00 ± 606.42	0.79 ± 0.41	2345 ± 680
Sad	0.83 ± 0.38	2452.57 ± 814.98	0.76 ± 0.43	2766 ± 970
Fearful	0.81 ± 0.40	2247.22 ± 775.56	0.55 ± 0.50	3063 ± 1269

Table 4.2 Baseline - Single Modality Accuracy and Reaction time

	Congruent Accuracy	Incongruent Accuracy	Paired t	Congruent RT	Incongruent RT	Paired t
Face	0.91± 0.11	0.74± 0.13	5.69***	2126± 399	2427± 504	-5.88***
Voice	0.81± 0.14	0.54± 0.16	8.20***	2392± 530	2690± 533	-4.66***
Happy	0.95 ± 0.11	0.68 ± 0.12	8.55***	1730 ± 374	2228 ± 413	-11.05***
Sad	0.81 ± 0.14	0.70 ± 0.15	3.98**	2586 ± 567	2583 ± 512	0.04
Fearful	0.82 ± 0.14	0.63 ± 0.19	5.81***	2462 ± 413	2681 ± 592	-2.73*
Overall	0.86±0.10	0.64 ± 0.10	11.50***	2259±425	2497 ± 481	-4.92***

Table 4.3 Accuracy of Congruent and Incongruent pairs

4.5.1.1 Facilitation effects

Accuracy (see Figure 4.1(a) and 4.1(b))

(a) (b)

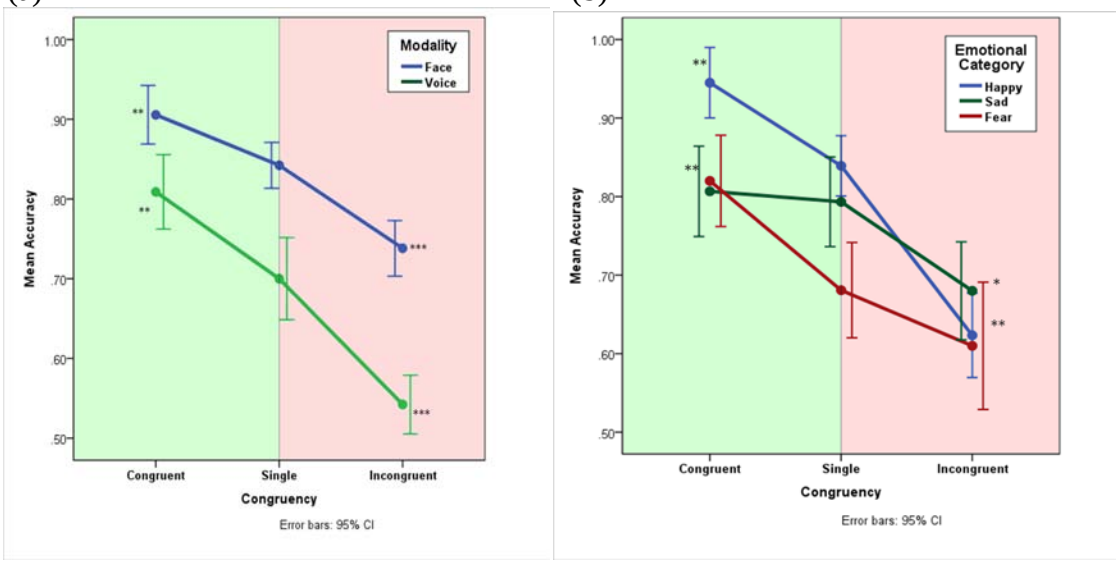


Figure 4.1 (a) Mean accuracy for Face and Voice experiments under different congruency state; (b) Mean accuracy for emotions under different congruency state

As expected, participants were more accurate on congruent trials than on single-modality trials ($F(1,24)$=16.35, p<0.001). Across the single and congruent dual-modality trials, accuracy was higher during the Face experiment (when faces were the target) than in the Voice experiment (when voices were the target) ($F(1,24)$=41.22, p<0.001, see Figure 4.1(a)). The main effect of emotion showed that accuracy was highest for happy, worst for fear and intermediate for sad target emotion trials ($F(2,48)$=17.82, p<0.001, see Figure 4.1(b)). This effect was qualified by a *single/multisensory*emotion* interaction (F(2,48)=5.03, p=0.01, see Figure 4.1(b)). Post-hoc analysis demonstrated greater accuracy facilitation (more positive *rAcc*) on happy and fear target emotion trials (*p<0.05 and*

p=0.05, respectively, see Figure 4.1(b)). All other interactions, including the 3-way interaction, were non-significant.

Reaction times (see Figure 4.2(a) and 4.2(b))

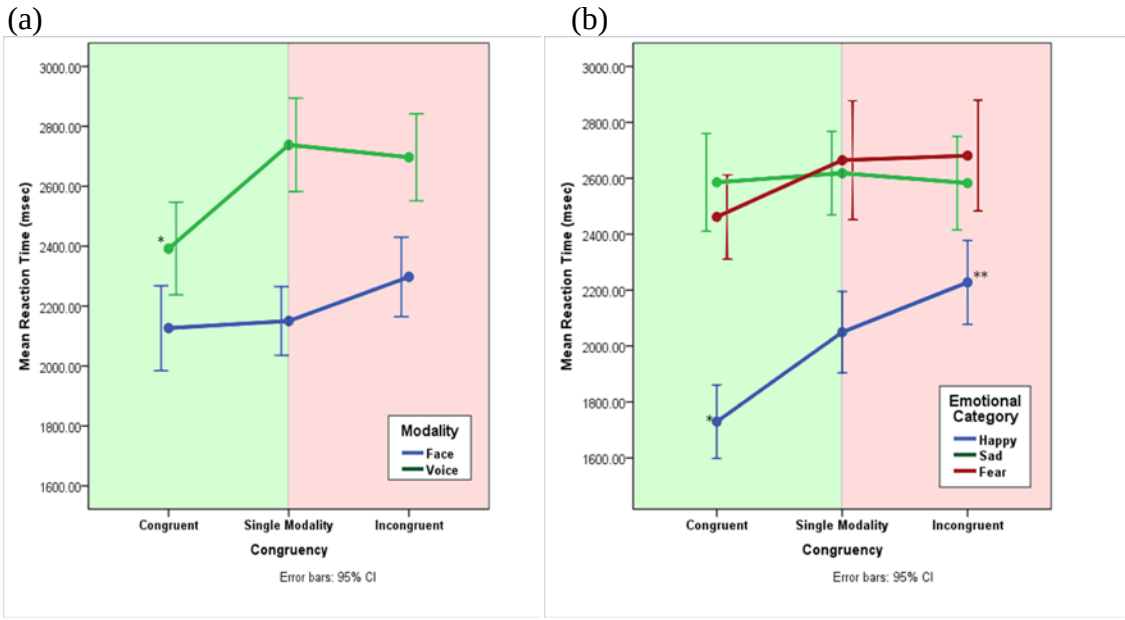


Figure 4.2 (a) Mean accuracy for Face and Voice experiments under different congruency state (b) Mean reaction time for emotions under different congruency state

As expected, participants responded faster on congruent trials than on single-modality trials ($F(1,24)$=7.44, p<0.05). Across both single and congruent dual-modality trials RTs were shorter during the face experiment than during the voice experiment ($F(1,24)$=41.40, p<0.001, see Figure 4.2(a). The main effect of emotion showed that RTs were shorter for happy relative to both sad and fear target emotion trials ($F(2,48)$=154.43, p<0.001, see Figure 4.2(b). This was qualified by a *single/multisensory*emotion* interaction ($F(2,48)$ =9.18, p<0.01, see Figure 4.2(b). Post-hoc analysis revealed greater RT facilitation (more negative *rRT*) on happy relative

to both sad and fear target emotion trials (*p<0.05 and p=0.05, respectively*, see Figure 4.2(b). There was also a *single/multisensory*modality* interaction (*F(1,24)* =12.08, $p<0.01$), driven by greater RT facilitation in the voice relative to the face experiment ($p<0.001$). All other interactions, including the 3-way interaction, were non-significant.

4.5.1.2 Interference effects

Accuracy (see Figure 4.1(a) and 4.1(b))

As expected, participants were less accurate on incongruent trials than on single-modality trials (*F(1,24)*=19.28, $p<0.001$). Across both single and incongruent dual-modality trials accuracy was higher during the face experiment than during the voice experiment (*F(1,24)*=73.79, $p<0.001$, see Figure 4.1(a). The main effect of emotion showed that accuracy was greatest for happy and sad relative to fear target emotion trials ($F(2,48)=13.74$, $p=0.001$, see Figure 4.1(b). There was also a *single/multisensory*modality* interaction (*F(1,24)*=6.49, $p<0.05$, see Figure 4.1(b), driven by greater accuracy interference (negative *rAcc*) in the voice relative to the face experiment (*p<0.05*, see Figure 4.1(a)). All other interactions, including the 3-way interaction, were non-significant.

Reaction times (see Figure 4.2(a) and 4.2(b))

Unexpectedly, RTs for incongruent trials did not differ from those for single modality trials (*F(1,24)*<1). Participants responded faster during the face experiment than during the voice experiment (*F(1,24)*=63.84, $p<0.001$, see Figure 4.2(a)). The main effect of

emotion showed that RTs were shorter for happy relative to both sad and fear target emotion trials (*F(2,48)*=102.33, $p<0.001$, see Figure 4.2(b)). This was qualified by a tend towards a *single/multisensory*emotion* interaction ($F(2,48)=3.53$, $p=0.08$, see Figure 4.2(b)). Post-hoc analysis revealed greater RT interference (more positive *rRT*) on happy relative to sad target emotion trials (*p<0.05*, see Figure 4.2(b), the latter of which in fact showed a slight numerical facilitation (negative *rRT*). There was also a trend towards a *single/multisensory*modality* interaction ($F(1,24)=3.61$, $p=0.07$), driven by RT interference (positive *rRT*) in the voice experiment, but no such interference (in fact a slight numerical facilitation: negative *rRT*) in the face experiment. All other interactions, including the 3-way interaction, were non-significant.

4.5.1.3 Correlations with schizotypy

In-line with our hypothesis, for congruent trials, there was a positive correlation between schizotypy and *rRT* ($r=0.42$, $p<0.05$), indicating that participants with higher levels of schizotypy benefited less from the facilitating effect of congruent stimuli. There was no association between levels of schizotypy and *rAcc* ($r=0.11$, $p>0.05$) (see Figure 4.3(a) and Table 4.4). However, contrary to our hypothesis, for incongruent trials there was also a positive correlation between schizotypy and *rRT* ($r=0.62$, $p<0.01$), indicating that participants with higher levels of schizotypy were slowed more by the interference effect of incongruent stimuli. Again, there was no association between schizotypy and *rAcc* ($r=-0.14$, $p>0.05$) (see Figure 4.3(b) and Table 4.4). Taken together these results indicate that compared to single modality trials, participants with higher levels of schizotypy spent more time processing dual modality emotional stimuli, without significantly altering their accuracy.

We also hypothesised that if negative processing bias was an influential process during emotional processing, participants with higher levels of schizotypy would have better accuracy and shorter reaction times for negative target emotions relative to positive target emotions (i.e. higher *dAcc[negative-positive]* and lower *dRT[negative-positive]*). These analyses were performed for both single modality and dual modality trials.

Congruency Influence Indexes	**Mean ± S.D.**	**SPQ**	**DASS**
rAcc	0.19 ± 0.09	-0.05	0.24
rAcc in congruent condition	0.09 ± 0.11	0.11	0.12
rAcc in incongruent condition	0.11 ± 0.12	-0.14	0.18
rRT	238 ± 242	0.41*	0.32
rRT in congruent condition	185 ± 339	-0.42*	-0.33
rRT in incongruent condition	53 ± 391	0.62**	0.48*

Table 4.4 Overall congruency influence indexes correlate with level of schizotypy and affective symptoms

The data did not support either hypothesis. In both single and dual modality tasks (averaging across congruent and incongruent trials), levels of schizotypy did not correlate significantly with *dAcc [negative-positive]* (r=0.15, p=0.47 and r=-0.07, p=0.75, respectively) and *dRT [negative-positive]* (r=-0.31, p=0.13 and r=-0.16, p=0.46 respectively).

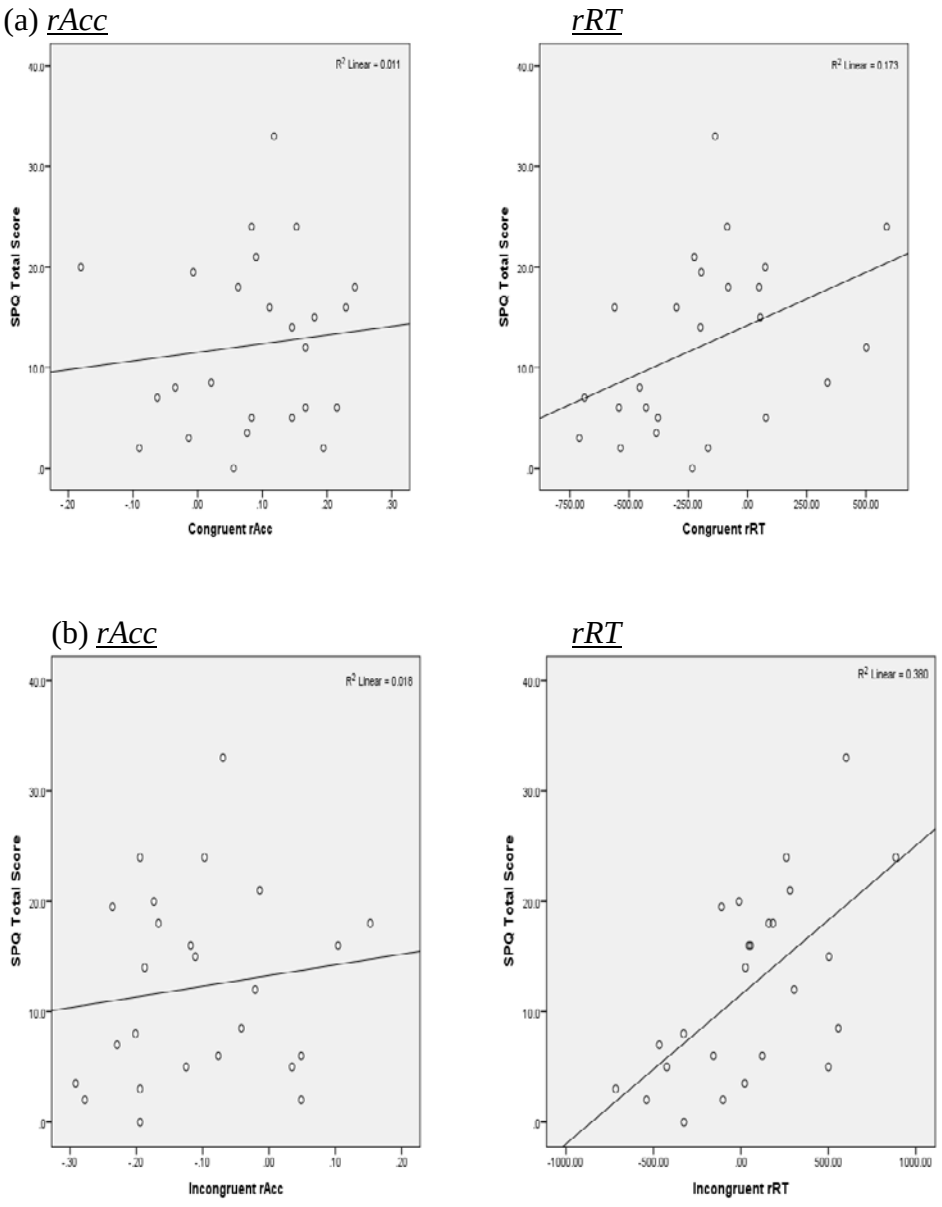


Figure 4.3 Scatter plots for correlations of (a) level of schizotypy and overall accuracy and RT for congruent stimuli across emotions (b) level of schizotypy and overall accuracy and RT for incongruent stimuli across emotions

4.6 Discussion

In the present study we tested a novel multisensory emotional recognition integration task (MERIT) which combined dynamic facial affect expressions with validated vocal prosodic emotions in healthy volunteers. Using the MERIT we demonstrated both facilitation and interference effects on emotional recognition accuracy and reaction times, relative to single modality trials, brought about by the presentation of congruent and incongruent visual and auditory emotional stimuli. As predicted, congruent emotional stimuli increased accuracy and reduced reaction times for emotional recognition, whilst incongruent stimuli decreased accuracy relative to single modality trials, although reaction times were not significantly increased. The largest facilitation and interference effects were seen for happy emotional stimuli. Significant facilitation and interference effects for accuracy were also seen for fearful emotional stimuli. However, no facilitation effects were observed for sad emotional stimuli, but a significant interference effect for accuracy was seen. We also examined the correlation between psychometric schizotypy and integration effects. As predicted, higher levels of schizotypy were associated with reduced facilitation effects in terms of longer RT during congruent emotional trials, though no association was found with accuracy. However, contrary to our prediction, during incongruent trials, higher levels of schizotypy were associated with increased interference effects in terms of longer RT although no association with accuracy was seen.

4.6.1 Emotional facilitation and interference effects

These results are broadly consistent with previous studies presenting simultaneous non-emotional (McGurk and MacDonald, 1976, Miller, 1982, Schröger and Widmann, 1998,

de Gelder and Vroomen, 2000) and emotional multisensory sensory stimuli (de Gelder et al., 1999, Dolan et al., 2001, Collignon et al., 2008, Jeong et al., 2011). Several studies demonstrated integration effects when subjects processed congruent and/or incongruent information across various emotions (e.g. happy, sad, disgust and fear). For example, tasks using static facial expressions presented with spoken neutral prosodic words/sentences(de Gelder and Vroomen, 2000), dynamic facial expressions with non-linguistic vocal clips of affect expressions (Collignon et al., 2008), and static emotional faces with emotional instrumental music (Jeong et al., 2011) all produce integration effects. These findings suggest that the emotional audiovisual integrative process in healthy subjects is as automatic in nature as the non-emotional one (Muller et al., 2011).

The overall non-significant incongruent interference effect on reaction times was inconsistent with previous studies (Ethofer et al., 2006b, Wittfoth et al., 2010, Muller et al., 2011). Intriguingly, the interference effect for RTs was significant for positive (happy) but not for negative emotions. This finding, together with the lack of facilitation effect for accuracy during sad emotional trials, suggests an emotional-specific effect for emotional multisensory integration. Emotional specificity may be due to separate neural networks used to integrate bimodal information (Park et al., 2010). However, psychometric properties of the emotional stimuli should be considered. Negative emotional faces and voices usually have less agreement, relative to positive emotional stimuli, in healthy volunteers (Tottenham et al., 2009) who take longer to respond to these stimuli(Leppanen et al., 2003, Hoekert et al., 2008). This effect might also be seen for bimodal emotional stimuli. Reaction times increase during dual modality trials where positive target stimuli are presented with non-target negative stimuli. However,

this increase is not seen during dual modality trials where negative stimuli are presented with non-target positive stimuli.

Across all trials, subjects identified the emotion of faces quicker and more accurately than voices. This difference was also seen during single modality trials. For morphing faces, visual emotional information is within a spatial domain but the intensity is incremental with time. However, for prosodic voices, auditory emotional information can only be obtained along the time domain. Together with greater attention-capturing nature of visual stimuli during audiovisual presentation (Koppen and Spence, 2007), the result is in line with a visual dominance effect in non-emotional bimodal perception(Colavita, 1974). Another possibility is that, during the MERIT, participants were instructed to make a judgment as soon as possible before the clip ended, meaning that they did not always listen to the whole clips of the voice prosody before making a judgment, reducing accuracy. Nevertheless, the *single/multisensory*modality* interaction for accuracy was non-significant, suggesting the difference did not influence the overall integration process.

4.6.2 Correlations with schizotypy

Previous studies of multisensory emotional integration report that patients with schizophrenia demonstrate decreased cross-modal facilitation effects relative to healthy control subjects (de Gelder et al., 2005, de Jong et al., 2009, de Jong et al., 2010). Based on a continuum model, we therefore hypothesised that higher levels of schizotypy in healthy volunteers would also be associated with altered multisensory integration,

resulting in reduced facilitation effects during congruent trials, and reduced interference effects during incongruent trials. In line with our hypothesis, for congruent trials, there was no association between levels of schizotypy and *rACC* whilst there was a positive correlation between schizotypy and *rRT*. This indicates that participants with higher levels of schizotypy benefited less from the facilitating effect of congruent stimuli on reaction times. However, contrary to our hypothesis, higher levels of schizotypy were also associated with greater *rRT* during incongruent trials, although there was no association with *rACC*. In other words, individuals with higher schizotypy scores showed *increased* interference effects for incongruent stimuli at least in terms of increased reaction times, although accuracy was unaffected.

Taken together, these results are unclear and do not allow us to accept our hypothesis that higher levels of psychometric schizotypy are associated with altered multisensory integration. Participants with higher levels of schizotypy took longer to make a judgment during multisensory trials, regardless of whether the information was congruent or incongruent. Longer RT could be due to a speed-accuracy trade-off (Grice and Spiker, 1979); however, we consider this interpretation unlikely as *rACC* scores did not correlate with schizotypy levels, for either congruent or incongruent trials. This pattern of results suggests an increased difficulty in processing information on all dual modality trials in individuals with higher schizotypy (i.e. extra information during dual modality trials impaired performance, regardless of congruency). This may be indicative of a more general cognitive deficit [46] which is exposed under conditions of high processing load and results in less effective utilisation of multisensory emotional information. However, the extent to which impaired utilisation of multisensory

information is secondary to or independent of a general cognitive deficit is unclear. Furthermore, this interpretation is not definitive as performance was only impaired in terms of RT and not accuracy.

We also hypothesised that greater schizotypy would be associated with a more negative processing bias. We did identify an association between schizotypy and affective symptoms (r=0.50, p<0.05), which is in line with an established literature reporting higher negative affectivity in schizotypy and schizophrenia spectrum disorders (Horan and Blanchard, 2003, Suslow et al., 2003, Horan et al., 2008, Blanchard et al., 2011, Najolia et al., 2011). However, contrary to our hypothesis, schizotypy was not associated with greater accuracy and quicker reaction time for negative target stimuli relative to positive target stimuli, during either single or dual modality trials. This suggests that negative biases did not affect the processing of single or dual modality emotional information in those with higher schizotypy.

As reviewed in Chapter 2, a number of confounding factors are inherent to studies investigating possible MSI deficits in schizophrenia, such as illness chronicity and the use of antipsychotic medications. In this regard, studies on patients with a first episode of psychosis, or in subjects with prodromal symptoms, provide an opportunity to circumvent some of these limitations. A number of studies in patients in the early and prodromal stage of psychosis show anatomical and functional abnormalities in neural circuitry associated with emotional processing (Seiferth et al., 2008; Modinos et al., 2010; Smieskova et al., 2010) relative to healthy controls, which may further affect the

MSI ability. Further investigation of the neural substrates of MSI in these cohorts is warranted to clarify this issue.

4.7 Limitations

The main limitation of the present study is the small sample size. Though large enough to demonstrate facilitation and interference effect, the sample size was not optimal to test for associations between schizotypy and multisensory integration. In particular, the negative relationship between *dRT* and schizotypy on single modality trials (r=-0.31), which is indicative of a more negative bias more schizotypal individuals, might reach significance in a larger sample.

4.8 Conclusion

Multisensory emotional integration is an important neuropsychological function that helps the individual organise emotional information from different sensory modalities, and in turn facilitates effective interactions with others. The current study demonstrates an association between psychometric schizotypy in healthy volunteers and slowed response during multisensory trials but not during single modality trials. This less effective utilisation of multisensory emotional information is unlikely to reflect impaired multisensory integration and may instead be due to a more general cognitive deficit under high processing load. Further research of multisensory emotional integration with a larger sample on participants with higher risk of psychosis is warranted to elucidate the role of emotional processing in the development of psychotic disorders. In the light

of the findings in the current study, a future design of the study is suggested to include an objective measure of attention function (for example, Cognitive Performance Task), and a non-emotional task with gradients of processing load (for example, N-back task) to confirm the interpretation of these findings.

Chapter 5

Dynamic Facial Emotional Expression Recognition in Early and prodromal Phase of Psychosis

5.1 Introduction

Dysfunction of emotional processing exists in first-episode psychosis (FEP) patients and might have already been present before the onset of clinical symptoms (Walker et al., 1993). These deficits are qualitatively similar but less severe to those seen in patients diagnosed with schizophrenia or other psychotic disorders (Jones et al., 1993, Yung and McGorry, 1996). They are significantly related to prodromal positive symptoms and general psychopathology in high-risk individuals (Eack et al., 2010), and might predict the subsequent onset of psychosis (Demjaha et al., 2010, Raballo et al., 2011). Hence, dysfunction of emotional processing may be an early precursor to positive symptoms and ultimately the development of schizophrenia. However, to date, research into neural alterations in the prodromal phase of psychosis has focused mainly on cognitive rather that emotional deficits (Fusar-Poli et al., 2007b).

Emotional perception and recognition is the initial stage of emotional processing and processing problems at the initial perceptual stage will have cascading impacts on higher order processing, resulting in altered emotional experience. Though emotional disturbances and affective symptoms are frequently presented in the UHR (Yung et al.,

2003a), they are not pathognomonic for schizophrenia or specific psychotic disorders. Investigating emotional perception and recognition, the basic and initial component of emotional processing in the UHR, may reveal distinctive and more specific changes of emotional disturbance to the UHR population. Difficulties with emotional perception in patients with schizophrenia have been widely studied in both the visual and auditory sensory modalities (Edwards et al., 2002). Deficits in facial emotion recognition have also been demonstrated in UHR populations (Addington et al., 2012, Amminger et al., 2012a, Amminger et al., 2012b, Thompson et al., 2012) and FEP patients (Pinkham et al., 2005, Reske et al., 2009, Brown and Cohen, 2010), with UHR subjects demonstrating an intermediate performance between FEP patients and controls (Thompson et al., 2012). This suggests that emotional recognition deficits in the UHR population are less prominent and that the performance of UHR subjects on these tasks may only differ from control subjects when subtle emotional cues or complex social cognitive paradigms are used (Thompson et al., 2012). Thus, in order to identify subtle emotional processing deficits in UHR subjects, and in the underlying neural correlates, a paradigm with higher ecological validity (i.e. using dynamic facial expressions instead of still pictures (Arsalidou et al., 2011), could be useful.

5.1.1 The neural correlates of facial emotional processing

Facial expressions are the nonverbal expressions most frequently encountered in daily life and deliver important emotional and social information. Processing of emotional faces involves discriminating the changes of geometric configuration of facial features based on the appearance (i.e. neutral faces) and recognition of the emotional meaning of

the pattern of change (Fusar-Poli et al., 2009). This process in real life is based on dynamic and continuous change of the facial geometric configuration. In the recent decade, the increasing interest in affective neuroscience has led to a large number of neuroimaging studies examining the fundamental process of social cognition, and the underlying neural mechanisms of facial emotional processing. A meta-analysis by Fusar-Poli et al. (2009) concludes that processing of emotional faces is associated with increased activation in a number areas, including 1) visual areas for early perception and decoding of facial features (inferior and middle occipital gyri, lingual gyrus and fusiform gyrus); 2) temporoparietal areas for integrating the static and movement features of face into a percept and roughly labelling it as emotional or not (parietal lobule, middle temporal gyrus, insula), 3) prefrontal areas for identification and fine categorisation (medial frontal gyrus), 4) limbic (including parahippocampal gyrus, posterior cingulate cortex) and subcortical areas (putamen) for associated knowledge for categorization and induction of emotional response within the subject, 5) the cerebellum.

Among these, neural responses in the visual cortex and cerebellum were observable across all emotional conditions; happy, fearful and sad faces specifically activated the amygdala, and insula was selectively activated during processing of disgusted and angry faces. In addition, age and gender also affect the neural response patterns during explicit or implicit processing of facial expression.

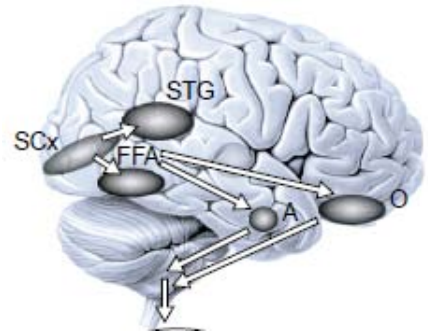


Figure 5.1 Emotional face processing (Adapted from Aldophs, 2002)

*STG = Superior Temporal Gyrus; FFA= Fusiform Area; A = Amygdala; O = Orbital frontal cortex; SCx = Striate cortex

Specific emotions seem to elicit distinct neural circuits. While compared with a baseline task condition (a fixation crosshair on the screen), the processing of neutral faces is associated with increased activation in limbic areas (left amygdala and left cingulate gyrus), subcortical areas (right lentiform nucleus), prefrontal regions and the left insula. When contrasted with neutral faces, happy faces are associated with neural activation in the bilateral amygdala, left fusiform gyrus and right anterior cingulate cortex, whereas sad faces activate the right amygdala and the left lingual gyrus. Fearful faces are associated with activation in the bilateral amygdala and the fusiform and medial frontal gyri (Fusar-Poli et al., 2009).

Recently a number of neuroimaging studies have employed tasks using dynamic facial expressions as these can elicit greater neurofunctional activity than static face stimuli. An Activation Likelihood Estimate (ALE) meta-analysis by Arsalidou et al. (2011) concluded that dynamic facial stimuli elicit increased activity relative to static faces in

regions associated with interpretation of social signals and emotional processing, including the middle temporal gyri and superior temporal sulci. Also, areas typically activated by facial emotion recognition tasks including amygdala, fusiform gyri, as well as the areas associated with cognitive manipulations, such as inferior frontal gyri are also activated during the processing and recognition of dynamic faces relative to a baseline condition (Arsalidou et al., 2011).

5.1.2 The neural correlates of emotional processing in schizophrenia and FEP

Neuroimaging studies demonstrate that schizophrenia is associated with neuroanatomical and neurofunctional abnormalities in the medial temporal lobe and the limbic structures including the amygdala, the hippocampus and the anterior cingulate (Gur et al., 2002a, Hempel et al., 2003, Williams et al., 2004). Functional alterations in these regions are thought to account for patients' characteristic disturbances in cognition and affective processing (Habel et al., 2004).

When processing emotional facial expressions, patients with schizophrenia show significantly less activation in the bilateral amygdala and right fusiform gyri. When compared to healthy control subjects, the extent of activation in bilateral amygdala, parahippocampal gyrus and fusiform gyrus, right superior frontal gyrus, and lentiform nucleus is significantly less in patients with schizophrenia (Li et al., 2010). Compared to implicitly perceived facial expressions, differences in activation of the fusiform area are more evident if the participants are asked to explicitly recognize facial expressions

(Li et al., 2010). This meta-analysis by Li et al. concluded that facial emotion processing difficulties could be due to an under-recruitment of the amygdala and substantial hypoactivation throughout the ventral temporal-basal ganglia-prefrontal cortex.

5.1.3 The neural correlates of emotional processing in UHR

Previous studies have shown that patients with FEP or UHR subjects differ from healthy subjects in terms of brain structure and function in many of the regions important for emotional processing (Fusar-Poli et al., 2007b, Smieskova et al., 2010). However, to date, few studies have explicitly focused on the neural substrate of abnormal emotional perception and other dysfunction of emotional processing in UHR subjects. A study conducted by Seiferth et al. (2008) showed increased activation in UHR subjects in the right lingual, the right fusiform, and the left middle occipital gyri relative to healthy controls during a facial emotion discrimination task. In another study, individuals who scored highly on a 'psychosis proneness' questionnaire had altered prefronto-limbic functional connectivity during emotion regulation relative to subjects with low psychosis proneness (Modinos et al., 2010).

Another region of interest in UHR subjects related to emotional processing is the valence-specific activation in the caudate nucleus. In healthy volunteers, the caudate nucleus has been shown to be sensitive to the negative emotional content of pictures

eliciting a greater response than for positive and neutral stimuli (Carretie et al., 2009). Interestingly, molecular imaging studies have provided *in vivo* evidence of increased dopamine synaptic availability and increased presynaptic dopamine synthesis in the striatum of people with UHR for psychotic illnesses (Howes et al., 2011), and functional alterations in the UHR groups have also been demonstrated in this region (Roiser et al., 2012).

5.1.4 Aims and Hypotheses

The aims of the study described in this chapter were to examine the neural substrate of facial emotion recognition in UHR and FEP groups relative to healthy controls using a dynamic facial recognition task. The hypotheses are as follows:

5.1) FEP subjects would show performance deficits in facial emotion recognition, relative to healthy controls, especially for negative emotional expressions (i.e. fear and sad faces). Associated alteration in neural substrate is expected in regions related to explicit facial emotion recognition, particularly bilateral fusiform area, left amygdale, and the right lentiform nucleus.

5.2) UHR subjects would show intermediate performance and activation deficits relative to healthy controls (i.e. the deficits, if present, would be less severe than those in FEP patients) during the dynamic faces task, particularly for negative emotions.

5.3) Given the role of the caudate nucleus in the processing of negative emotional stimuli, and the reported neurochemical and functional alterations in this region in UHR

groups, the UHR group will show altered activation in the striatum/caudate nucleus relative to controls during negative compared to positive emotional faces.

5.2 Materials and Methods

5.2.1 Participants

Participants' demographics, clinical data, and estimated FIQ are presented in Table 5.1. The study obtained ethical approval from the National Health Service UK Research Ethics Committee (reference number: 11/LO/0623). All subjects included in the study were 18-35 years old except one of the FEP subjects, and all spoke fluent English.

Exclusion criteria included a history of neurological disorder, prior head trauma resulting in loss of consciousness and/or hospitalisation, or any contraindications to exposure to a magnetic field (e.g. metal implants, or pregnancy). Any participants reporting excessive use of alcohol (>21 units per week for men and >14 units per week for women) or recent recreational drug use (use of cannabis, stimulants, hallucinogens, or opiates in the 2 weeks prior to the fMRI scan) were excluded.

To date, only few functional MRI studies in UHR groups have been conducted using an emotional processing task. Nevertheless, fMRI studies of multisensory integration in healthy participants and patient with schizophrenia suggest that group differences in activation (i.e. BOLD signal) are detectable in superior temporal regions, one of the a-priori region of interest in the current study, with sample sizes around 12 - 15 in each

group (Park et al., 2010; Szycik et al., 2009; effect size=2.57, estimated sample size >= 5, power = 0.80 and alpha=0.05). Functional MRI studies of emotional processing in patients with schizophrenia also suggest that group differences in activation are detectable in cortical and subcortical regions with similar sample sizes in each group (For example, Gur et al., 2002b, Kosaka et al., 2002, Habel et al., 2004). However, taking into account that UHR groups tend to exhibit subtle functional changes that are are less severe that those seen in established psychosis/schizophrenia, a larger sample size may be need to detect activation differences.

5.2.1.1 First Episode Psychosis

Eighteen subjects with a first episode of psychosis (see Section 1.1.2, page 19 for more detail) according to the ICD-10 criteria were recruited to the study through South London and Maudsley Early intervention teams (LEO, STEP, LEIS and COAST) (http://www.slam.nhs.uk). First, suitable FEP participants were identified via consultation with NHS consultants, registrars and other staff at early intervention teams. Potential FEP participants were then approached and the details of the project explained in full.

5.2.1.2 Ultra High risk (UHR) Subjects

Sixteen UHR subjects were recruited from OASIS (Outreach and Support in Southeast London), a clinical service for young people at high-risk of developing psychosis (Broome et al., 2005). The UHR state is defined according to the Personal Assessment and Crisis Evaluation (PACE) criteria (Yung et al., 1998) and confirmed via a detailed

clinical assessment using the Comprehensive Assessment of At Risk Mental States (CAARMS) (Yung et al., 2008a) (see section 1.1.1, page 17 for more detail). UHR participants meet at least one of the following criteria; a) attenuated psychotic symptoms, b) brief limited intermittent psychosis, or c) a significant decline in cognitive and social functioning over the past year, together with either schizotypal personality disorder or a first degree relative with a psychotic disorder. UHR participants were between the ages of 18 and 35 years old as this age range confers the highest risk for onset of psychosis.

5.2.1.3 Healthy Controls

Twenty-one gender-matched healthy control subjects (HC) were recruited via advertisements from the same geographical areas as the UHR/FEP participants (i.e. South and South East London Area). No control subjects met criteria for a DSM-IV-TR psychiatric disorder, fulfilled the PACE criteria for prodromal symptoms, nor had a first-degree family history of psychiatric disorders.

5.2.2 Data Acquisition

5.2.2.1 Neurocognitive Assessment and Estimate of Premorbid IQ

Each participant underwent a neurocognitive assessment consisting of four subscales of the Wechsler Adult Intelligent Scale III (WAIS III), including Digit Symbol Coding, Arithmetic, Block Design and Information Subtest. The four subtests version of the WAIS III is both an evaluation of general cognitive ability (full-scale IQ) and also a

brief evaluation of different neurocognitive domains including attention, executive function, visual spatial function and psychomotor speed, which might be potential confounders affecting participants' performance on the task.

5.2.2.2 Clinical Symptom Profile

The positive and negative syndrome scale (PANSS) (Kay, 1987) was administered to FEP and UHR subjects by a trained researcher. The Clinical Assessment of At Risk Mental State (CAARMS) (Yung et al., 2005) was administered and scored for each UHR and HC participant by a trained researcher. The CAARMS was not administered to FEP patients as it is designed for detecting subtle psychopathology in the prodromal state and may have ceiling effects in this group. These scores were used for FEP and UHR subjects as a measure of concurrent psychopathology, and for HCs as a confirmation that they were not experiencing any attenuated psychotic symptoms.

5.2.2.3 Self-report questionnaires

Both affective and social anxiety symptoms might be contributing factors influencing the accuracy of emotion perception. The perceptual difficulties in UHR and FEP participants are hypothesized to be, at least partially, independent to these affective symptoms. Thus, the following tools were used to quantify the symptom profile and current psychological state for each participant: Depression and Anxiety Symptom Scale (DASS-42) (Lovibond and Lovibond, 1995) was used to evaluate general mood symptoms including level of depressive mood and anxiety and the Liebowitz Social Anxiety Scale (LSAS) (Liebowitz, 1987) was used to characterise social anxiety.

5.2.2.4 Magnetic Resonance Imaging

Functional MRI (fMRI) is a non-invasive imaging technique for mapping brain activity spatially (Ogawa et al., 1990). It detects the differences of magnetic properties between oxygenated and de-oxygenated haemoglobin in red blood cells. A natural blood oxygen level dependent (BOLD) contrast in the MR signal is created in the tissues closest to blood vessels (highly oxygenated) relative to the surrounding tissues (less oxygenated) (Ogawa and Lee, 1990). When neuronal activity increases, blood circulation also increases as a compensatory consequence of oxygen consumption. The dynamic changes of blood flow during this neurovascular coupling are referred to as the haemodynamic response. The BOLD contrast is sensitive to this haemodynamic response and used as a proximal measure of neural activity over time. fMRI thus can be used to map the changes of haemodynamic response with time in each area of the brain during a task. This allows researchers to make inferences of brain areas relating to mental processes of interest.

All neuroimaging was conducted using a 1.5T MRI scanner (Sigma, LX-GE, Milwaukee, USA) at the Maudsley Hospital, London. The emotional recognition paradigm for fMRI used the single modality emotion recognition task stimuli adapted from the behavioural task reported in Chapter 4, (i.e. MERIT task). The design of the paradigm was exactly the same as the single modality task of behavioural experiment described in Chapter 4, except that a blank screen replaced the stimulus after each button press until the full length of stimulus had passed, (note – in the behavioural task

the next stimulus was presented immediately). This modification was aiming to provide consistent total scanning time across participants. The same modification was also applied on the dual modality task used in Chapter 7. To summarise, there was 96 facial trials and 96 voice trials, with mean length of 4.2 seconds for each trial. A one-second inter-stimuli interval with a fixation cross in the centre of the screen followed each stimulus. During the single-modality emotion recognition task, face and voice trials were pooled together and presented in a pseudo-random order and arranged into two sessions. During the task, subjects' responses (accuracy and reaction time) were recorded for subsequent performance analysis.

Functional images were acquired using a TR of 3000 ms, a TE of 40 ms, a flip angle of 90°, a slice thickness of 2.5mm with 0.5mm gap, a field of view of 24cm^2 and a 64x64 matrix. In total, 46 axial slices in parallel to the anterior commissure–posterior commissure (AC-PC) line were collected for each subject. Overall 427 image volumes were generated in the single modality task for each participant. The task takes 21 minutes (10 and 11 minutes for the first and second session, respectively)

5.2.3 Data Analysis

5.2.3.1 Demographic, neurocognitive and clinical measurement Data

Analyses of variance (ANOVA) and Chi-square tests were conducted using the IBM SPSS 19 statistical software package in order to compare all demographic, neurocognitive and clinical measurements (http://www-01.ibm.com/software/analytics/

spss/products/statistics/). Post-hoc paired t-tests with Bonferroni correction for three comparisons (i.e. comparisons for any two of the three diagnostic categories) were used to clarify the significant differences. Inferences were made at $p<0.05$.

5.2.3.2 Dynamic Facial Emotion Recognition: Behavioural Analysis

Based on digital recordings of subjects' responses during the task, the mean proportion of errors and mean reaction times were calculated for each subject. For the HC, UHR and FEB groups, the numbers of subjects completing the experiment were 20, 16 and 18, respectively. One healthy subject was excluded due to incomplete data collection.

Accuracy and RT were analysed using separate repeated measures analyses of variance (RM-ANOVAs). The emotional category (happy, sad, fear, neutral) was entered as the within-subject variable. Diagnostic group was entered as the between subject variable. Equal variances were assumed unless Levene's test of equality of variance was significant at the 95% level. Post-hoc paired t-tests with Bonferroni correction for three comparisons (i.e. comparisons for any two of the three diagnostic categories) were used to clarify the significant differences. Separate ANOVA for each emotion were then performed to further examine the group differences in individual emotion. Where there was a significant emotion-by-group interaction, a post-hoc pair-wise comparison with Bonferroni correction was performed for individual emotion conditions separately, with the accuracy and RT as the dependent variable and group as the fixed-factor.

5.2.3.3 Functional MRI

Functional images were pre-processed using SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) running under Matlab 7.1 (Math Works, Natick, MA, USA). Two single modality sessions were entered as two separate sessions, and a standard event-related design was conducted whereby the first images of both sessions were realigned to the obtained structural image. The remaining images were then realigned to the first image of their respective session and resliced with sinc interpolation.

Movement parameters were calculated and the image with excessive movement (>1.5 mm of translation and 1 degree of rotation in any axis) and the adjacent images were examined and removed if the image was corrupted. Interpolation of the images adjacent to the corrupted images was performed to replace the removed images. One UHR subject and one FEP subject were excluded during the analysis of functional MRI data due to excessive movement.

Finally, the images were segmented and spatially normalized (Friston et al., 1995) to a standard MNI-305 template using nonlinear-basis functions and spatially smoothed with a 8-mm full width at half maximum isotropic Gaussian kernel to compensate for residual variability in functional anatomy after spatial normalization, permitting the application of Gaussian random field theory for adjusted statistical inference. The groups entered into first and second level analysis comprised 20 HC, 15 UHR and 17 FEP subjects.

A standard event-related first-level analysis of regional responses was performed to identify regional activations in each subject; this involved convolving the onset times (i.e. the onset of the facial expression clips) with a canonical haemodynamic response function. To exclude low frequency drifts the data was high-pass filtered using a set of discrete cosine basis functions with a cutoff of 128sec. Ten experimental conditions were used: 1) Happy Face 2) Sad Face 3) Fearful Face 4) Neutral Face 5) High-intensity Happy Voice 6) Low-intensity Happy Voice 7) High-intensity Sad Voice 8) Low-intensity Sad Voice 9) High-intensity Fearful Voice 10) Low-intensity Fearful Voice. The movement parameters were entered as separate regressors of no-interest, as well as each corrupted scan which has been removed and replaced. Only analyses based on the face conditions are reported in this chapter. To remain consistent with the format of the data collected in the dual-modality experiment, error responses were not removed through the use of a separate error regressor; however, this raises the possibility of potential 'noise' in the data due to individual differences in behavioural performance which may in turn have reduced statistical sensitivity. Using the GLM parameter estimates obtained for all brain voxels for five 1st level contrasts of interest were then computed, Happy–Neutral, Sad–Neutral, Fear–Neutral, overall emotional stimuli–Neutral, and positive emotion (Happy) – negative emotion (Fear + Sad). Neutral faces served as the baseline contrast for each emotion.

Second level analysis of the contrast images was performed using an independent ANCOVA F-test or independent t-test implemented in SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) running under Matlab 7.1 (Math Works, Natick, MA,

USA), with age entered as a covariate and statistical inferences made at $p<0.05$ after FWE correction for multiple comparisons at cluster level.

In order to correct for multiple comparisons, four areas reported in a meta-analysis (Li et al., 2010) of facial emotion recognition studies between healthy controls and patients with schizophrenia were used. These were the bilateral fusiform gyrus, left amygdala and right lentiform gyrus. Additionally, the caudate nucleus (Carretie et al., 2009), was also chosen as a region of interest for testing hypothesis 5.3 - relating to valence-specific interaction effects during the negative emotions faces relative to positive emotional faces.

A small volume correction (SVC) with a sphere of 12 mm radius was used according to the coordinates of previous studies, including left fusiform gyrus (-39, -65, -13), Right fusiform gyrus (40,-52,-14), left amygdala (-21, -7, -8), right lentiform gyrus (22, -3, -5), left caudate body (-18, -2, 24) and right caudate body (16, 4, 18). All the coordinates reported were described using a standard Montreal Neurologic Institute (MNI) coordinate system.

5.2.3.4 Correlational Analysis

Pearson's correlation analyses were performed to examine the relationship between the level of activation and clinical symptoms, including PANSS score in both UHR and

FEP groups and CAARMS severity score in the UHR group. Subscales of positive symptoms and negative symptoms were also examined.

5.3 Results

5.3.1 Demographics and psychopathology

The demographic data and psychopathology for each group were reported in Table 5.1.

	HC (n=21)	UHR (n=16)	FEP (n=18)	$F(\chi)$	p
Age (years)	23.29±4.55	24.19±4.22	27.72±5.36	4.58	0.05 FEP>NC=UHR
Gender[c]	8M:13F	8M:8F	13M:5F	4.60	0.10
Laterality[c]	22R:0L	15R:1L	17R:1L	3.09	0.54
Years of Education	16.90±2.30	14.81±1.83	14.78±3.98	3.54	0.04 NC>UHR=FEP
Current FIQ	106.33±21.94	96.56±17.41	98.56±21.13	1.22	0.30
Ethnicity[c]	14W:4B:1A:2O	8W:8B:0A:0O	10W:6B:0A:2O	9.47	0.14
PANSS total[a]	-	52.80 ±11.33	54.56 ±13.79	0.16	0.70
PANSS positive[a]	-	12.60±2.92	13.47±5.29	1.36	0.25
PANSS negative[a]	-	14.39±6.24	13.17±5.45	0.02	0.88
PANSS general[a]	-	26.73±5.35	27.00±7.36	0.14	0.91
CAARMS total	2.33±3.81	34.27±17.35	-	67.33	<0.001
CAARMS positive	0.57±1.08	6.94±4.57	-	38.33	<0.001
CAARMS negative	0.05±0.22	4.31±3.38	-	33.54	<0.001

M = males; F = females; R = Predominantly Right Handed; L = Predominantly Left Handed; WRAT-R (SS) = Wide Range Achievement Test Revised (Standardized Score); PANSS = Positive and Negative Syndrome Scale; Ethnic Groupings - W = White Caucasian, B = Black African or Black Caribbean, A = Asian, O = Mixed/Others.

Table 5.1 Demographic information for the subjects used for each of the three diagnostic group comparisons, showing the mean followed by the standard deviation (in brackets)

5.3.2 MERIT facial recognition performance

Accuracy: A repeated measures ANOVA showed a significant main effect of emotion (F =10.30, df =3, p<0.001). Across subjects accuracy was highest for happy, lowest for sad and intermediate for fear. The main effect of diagnostic group was also significant (F =4.19, df =2, p=0.02) with highest accuracy in HC and lowest accuracy in FEP patients. The diagnostic group*emotion was non-significant. The accuracy for individual emotions are illustrated in Figure 5.2(a).

Reaction Times: A repeated measures ANOVA showed a significant main effect of emotion (F =77.69, df =3, p <0.001). The main effect of emotion was driven by shorter reaction times for happy emotional faces. Reaction times were longest for neutral and intermediate for sad and fear. Both the main effects of diagnostic group and diagnostic group*emotion interaction were non-significant (F =1.50, df =2, p =0.23). The reaction times for individual emotion are illustrated in Figure 5.2(b).

Figure 5.2(a)

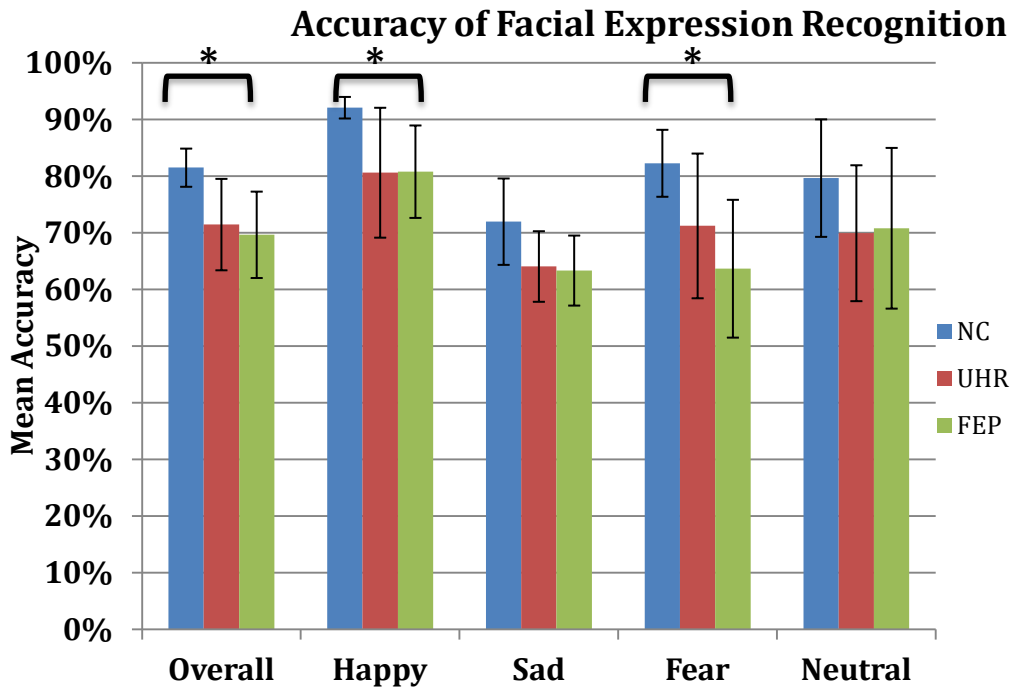


Figure 5.2(b)

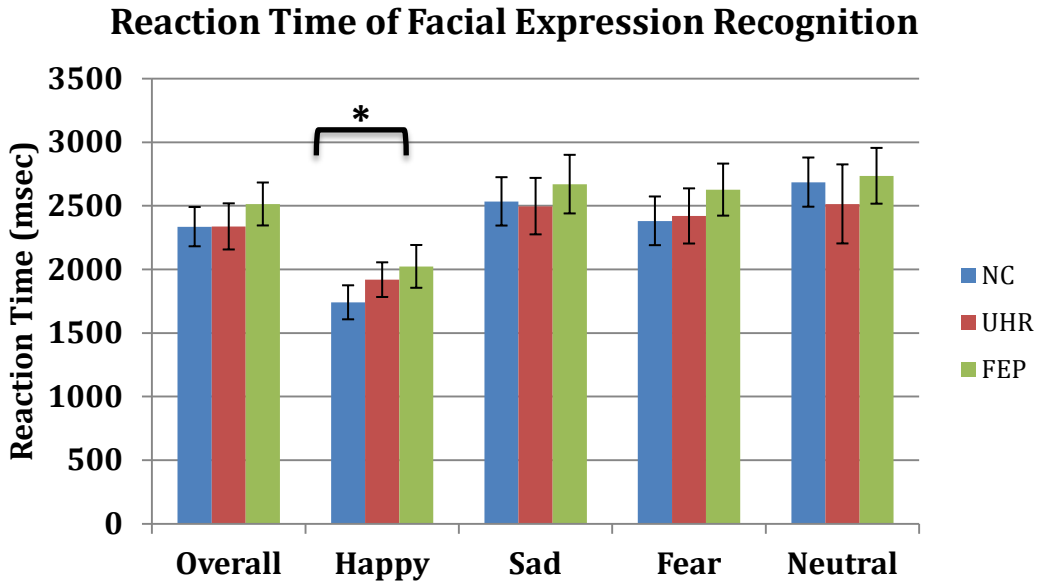


Figure 5.2 Graph showing (a) mean accuracy, and (b) mean reaction time for each group by emotional category.

5.3.3 Functional MRI

5.3.3.1 MAIN EFFECT OF TASK

Emotional voice > Near-neutral faces

Using the whole brain analysis a main effect of emotion across groups was seen in the bilateral lingual and fusiform gyrus, extending to precuneus and left posterior cingulate gyrus, as well as left middle temporal and occipital gyrus, extended to pSTS ($p<0.05$, FWE corrected with cluster size >493 voxels; see Figure 5.3).

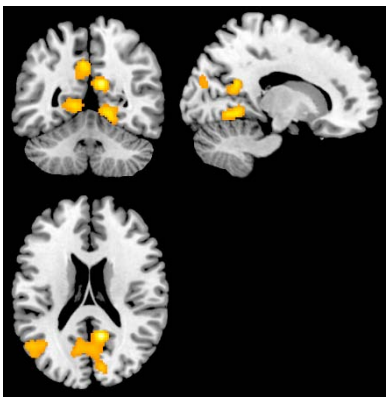


Figure 5.3 Peak BOLD activation for facial emotion task effect across groups

Positive versus Negative faces

Using the whole brain analysis, an effect of positive > negative emotion across groups was seen in the both sides of the angular gyrus, bilateral precuneus, posterior cingulate gyrus, left parahippocampal gyrus and inferior orbital frontal and anterior cingulate gyrus ($p<0.05$, FWE corrected). An effect of negative > positive emotion across groups

was seen in left precentral gyrus, supplementary motor area, bilateral middle occipital lobe, and bilateral lingual/calcarine area (p<0.05, FWE corrected; see Figure 5.4).

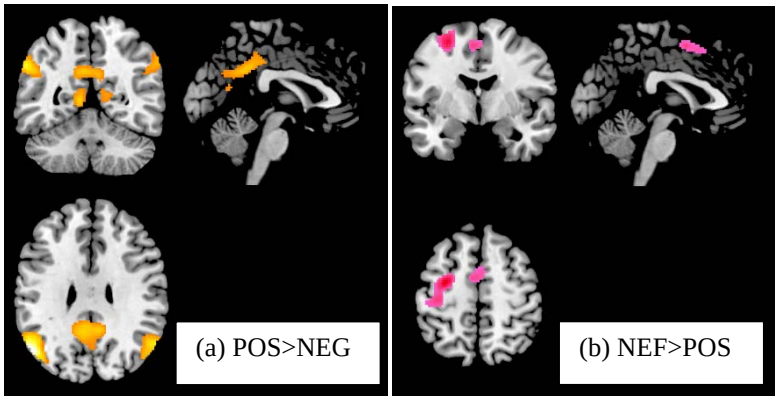


Figure 5.4 Peak BOLD activation for emotion effect across groups. (a) positive emotion > negative emotion (b) negative emotion > positive emotion

5.3.3.2 MAIN EFFECT OF GROUP

Using the whole brain analysis the main effect of group was non-significant.

5.3.3.3 GROUP × TASK INTERACTION

Using the whole brain analysis the group × task interaction was non-significant.

5.3.3.4 REGION OF INTEREST ANALYSIS

Using the ROIs and coordinates described above, an average effect of tasks (happy, sad, fearful faces versus neutral faces) was seen in the left amygdala (p=0.03, t=3.51). A group main effect was seen in the right fusiform gyrus (p=0.001, F=9.75). Relative to HC; in both FEP patients (p=0.002, t=3.91) and UHR groups (p=0.009, t=3.23) decreased activation for emotional faces relative to neutral faces was seen in right

fusiform gyrus (see Figure 5.5). The group differences in the other regions were non-significant. The group by emotion interaction was not significant.

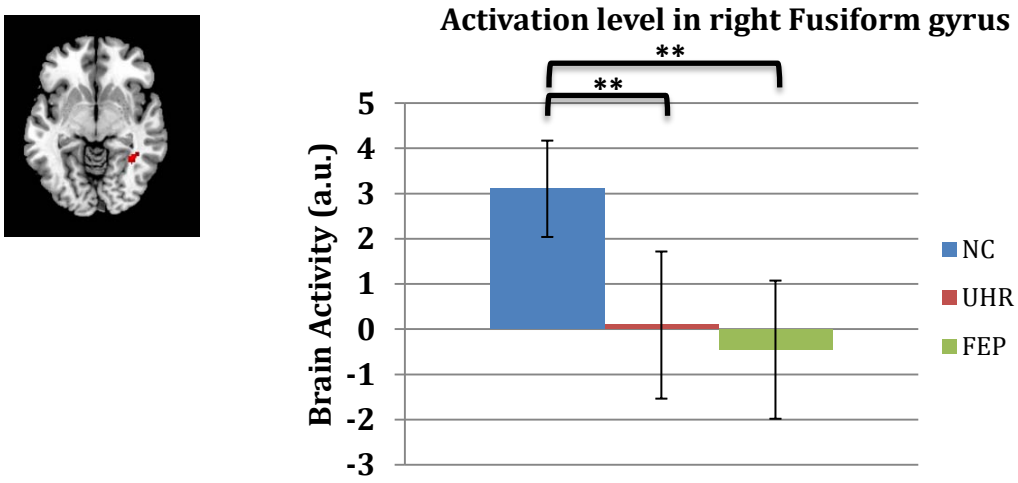


Figure 5.5 Graph showing peak BOLD activation level in right fusiform gyrus for each group during emotional faces contrasting to neutral faces.

5.3.4 Correlation Analysis

Using the peak brain activity extracted from the region of interests showed significant group effects (i.e. the right fusiform gyrus at the current analysis), correlation analyses with behavioural performance and clinical symptoms were performed. There were no significant correlations between activation levels and performance, including accuracy, duration and number of missing trials. There were no significant associations between the activation levels and PANSS scores nor any subscales in both UHR and FEP group. There were significant negative correlations between peak activation in the right fusiform gyrus and levels of emotional symptoms on the CAARMS ($r=-0.47$, $p = 0.004$), across NC and UHR group.

5.4 Discussion

The current study aimed to investigate whether the neural substrate of emotional processing was altered in UHR and FEP groups relative to a HC group using a dynamic emotional faces task.

5.4.1 Behavioural Results

In line with hypothesis 5.1, the FEP group was less accurate than the HC group in recognising dynamic emotional facial expressions, particularly during fear and happy trials. As expected in hypothesis 5.2, the performance of the UHR group was numerically intermediate between the HC group and the FEP group, although the difference between UHR and the other two groups was not statistically significant. Reaction times were significantly slower in FEP for happy trials only, and overall all three groups responded with comparable speed. These results are consistent with the finding from Amminger et al. (2012a) who report that impaired facial emotion recognition may be apparent before the full expression of psychotic illness in UHR subjects. Although impairment in UHR subjects is not as severe as that seen in first episode psychosis, performance may not be differentiable from the latter group (Thompson et al., 2012).

5.4.2 Whole brain analysis

The main effect of emotion across groups was seen in the bilateral lingual and fusiform gyrus, extending to precuneus and left posterior cingulate gyrus, as well as left middle temporal and occipital gyrus, extended to posterior STS. The finding in the temporal-

parietal visual area is consistent with the literature (Adolphs, 2002, Fusar-Poli et al., 2009). The activation in the limbic system and frontal area could also be seen without correction ($p<0.001$), but did not survive after correction for multiple comparison. Contrary to hypotheses 5.1 and 5.2, despite robust task effects across groups, no significant whole brain (voxel-wise) group effect was observed.

5.4.3 ROI Results

The main effect of emotion across groups in the left amygdala can be seen in the ROI analysis. The finding is consistent with the model of emotional recognition, reflecting the amygdala involvement during the identification/interpretation stage and emotional appraisal (Adolphs, 2002, 2008, Fusar-Poli et al., 2009).

In line with hypotheses 5.1 and 5.2, using an ROI approach both FEPs and UHRs had significantly decreased activation in right fusiform gyrus, and decreased activation in the UHR group was intermediate between the HC and FEP group. Such difference remained significant after including age as covariates in the second-level analysis. Such finding is broadly consistent with previous studies showing structural and functional abnormalities in the fusiform gyrus in schizophrenia (Quintana et al., 2003, Marwick and Hall, 2008), and first-episode schizophrenia (Lee et al., 2002).

Previous studies have shown altered structure and function in UHR cohorts across multiple cortical and subcortical brain regions (Borgwardt et al., 2007, Egerton et al., 2011). Although few studies have focused on emotion and social cognition, our results are partially consistent with Seiferth et al. (2008) who reports decreased activation in right lingual and fusiform gyrus during a facial emotion discrimination task relative to HCs.

Nevertheless, the level of activation was not associated with any measure of task performance across groups or within each group. Furthermore, the group by task (positive vs. negative emotions) interaction was not significant, suggesting that reduced activation in the fusiform face area during the dynamic facial emotion expression task is not emotion-specific, but rather a generic effect found across positive and negative emotions in both FEP and UHR groups. A configural face processing deficit in UHR and FEP during facial emotional processing, which is similar to that reported in schizophrenia (Shin et al., 2008), rather than a specific emotional processing deficit, could have accounted parsimoniously for these findings. However, impairments in facial emotion recognition may not be exclusively secondary to impairments in face recognition, as the facial emotion recognition deficits remains significant after statistically controlling for facial identity recognition impairment (Penn et al., 2000).

Contrary to hypothesis 5.3, the group differences in activation of the striatum and caudate area were non-significant in both whole brain analysis and ROI analysis. The group differences in activation of striatum and caudate area were non-significant, which

might be related to the following: Firstly, although there are neurofunctional differences between UHR and HC in the striatal area (Howes et al., 2011), the current emotional recognition task might not be optimal to demonstrate the differences. Anger and disgust are the two negative emotions most significantly showing the activity in these areas in healthy controls (Fusar-Poli et al., 2009), but they have not been included in the design due to time limitations on the scanning session and less significant behavioural differences reported between NC and schizophrenia (Edwards et al., 2002). Secondly, to demonstrate the emotion-specific effect of facial emotion processing on activation, neutral emotion was chosen as a baseline. However, neutral facial emotion activates the subcortical striatum area in healthy participants (Fusar-Poli et al., 2009). The existing differences in this area might have partially been removed before entering the analyses contrasting positive and negative emotions. Finally, although valence-specific activation pattern exists in the healthy group (Carretie et al., 2009), the possibility still exists that there might not be significant neurofunctional differences specific to emotional valence among the three groups during the facial emotional processing.

5.4.4 General discussions

Taken together, the results of the present study show that FEP and UHR subjects performed significantly worse than HCs during an emotional recognition task and had significant reduction in activation of the fusiform gyrus, consistent with the literature suggesting that deficits in emotional processing are another important feature of both established psychosis and UHR states (Amminger et al., 2012a, Thompson et al., 2012).

Furthermore, the results also show that impaired emotional recognition, as well as the reduced activation in right fusiform gyrus was similar in FEP and UHR groups, giving some support to the notion that facial emotional processing deficits may occur early in the course of psychosis. The deficits might be associated with decoding facial information into emotional content, rather than emotional processing per se.

The finding that there were no differences in amygdala activation warrants further discussion. Firstly, though there is consistent evidence suggesting amygdala dysfunction in patients with schizophrenia (Lee et al., 2002, Li et al., 2010, Tian et al., 2011), amygdala dysfunction in FEP and UHR subjects has not been established. It is possible that amygdala dysfunction in FEPs and UHR states may not be as prominent as it is in established schizophrenia, or amygdala dysfunction is only associated with fully blown psychotic states when permanent neural processing changes have been established.

The finding that decreased activation can be found only in the fusiform gyrus but not amygdala implies a differential neurofunctional impairment in the early stage of psychosis. Early neurofunctional alteration, before the full-blown onset of psychosis, in the fusiform gyrus may potentially be a trait-related neurofunctional biomarker. This area is responsible for the configuration of faces during emotional processing, and the alteration may be associated with both early visual perceptual disturbances, and dysfunctional emotional information processing. In contrast, functional alterations of amygdala are more state-related changes associated with fully blown psychotic states when permanent neural processing changes have been established.

Alternatively, the amygdala dysfunction may exist in some individuals in the early or prodromal stage of psychosis, but the clinical manifestations in FEP and UHR cohorts were much more heterogeneous compared to the schizophrenia cohorts that have been studied in previous neuroimaging investigations of amygdala dysfunction. Higher levels of inter-subject variability may have decreased the statistical power to detect true group effects (i.e. type II errors). The inherent heterogeneity is especially relevant when sample size was comparatively small, although our sample size was in the optimal range for detecting large effects while minimizing the risk of picking up trivial effects in a more homogenous cohort (Friston, 2012).

Methodological issues in analysing the imaging data might also have contributed to the non-significant neurofunctional results of emotion across groups. In the current analysis, errors trials were still included without being modelled as a separate regressor for two reasons: 1) even when the emotion had been misidentified, the participants were still processing emotional information, and the differences during the processing might reflect the possible alteration in neural activity among groups. 2) Due to the lower recognition rate in the FEP group, removal of error trials might significantly decrease the statistical power for detection the group difference. As a result of including error responses, the variation during the error responses might increase the noise and reduce the chance of detecting a reliable group level effect. However, although the removal of error trials could better show the neural correlates of individual emotions *per se,* the alteration related to misjudgement of emotion could be missing.

5.5 Limitations

The main limitation of the present study is the relatively small sample size. The size was sufficient to demonstrate the alteration of neurofunction in the FEP, but might not be optimal to fully demonstrate the subtle change in the UHR, in particular, the emotion-specific effect. The fact that the majority of FEP subjects were older in age could be another limitation of the study. Whilst the age difference was small and might not have an impact on accuracy, it remains a potential confound with respect to the findings observed here. Secondly, a longitudinal perspective will be helpful to delineate whether the changes are state or trait related, but the cross sectional design of the current study was not able to provide the information of changes over time. Finally, it should be noted that the dynamic facial emotion recognition is a relatively cognitive-demanding task to perform under the time constraint, and as such, for those subjects with a higher degree of cognitive and functional decline their performance could be confounded with their general cognitive ability instead of showing a pure effect of emotional processing. Nevertheless, regarding social cognition, it is not possible to completely exclude the influence of cognitive function; the results may still remain of value under reasonable interpretation for the alteration of social function in UHR and FEP.

5.6 Conclusion

To conclude, consistent with the majority of the literature, the findings suggest that FEP subjects were significantly worse than HC in facial emotion recognition, and the UHR subjects performed intermediately between HC and FEP subjects. Consistent with previous studies, the analysis was not able to detect significant differences in UHR relative to FEP subjects at the behavioural level. The analyses of neural activity during dynamic facial emotion recognition showed a significantly lower activation level in right fusiform gyrus in both UHR and FEP subjects. The activation level in the UHR was further correlated with higher positive symptoms and a lower level of social functioning.

Chapter 6

Prosodic Emotional Recognition in Early and Prodromal Phase of Psychosis

6.1 Introduction

Prosodic emotion perception deficits are another important aspect of emotional processing dysfunction in schizophrenia. It refers to the evaluation and judgment for the non-lexical information within the prosody of the spoken sentences, which delivers emotional state and intention of the speaker (Edwards et al., 2002, van Rijn et al., 2005). Inability to accurately identify the emotional information within prosodic voice, mainly for negative emotions, has been established in schizophrenia with a fairly large effect size (*d*=-1.24, See meta-analysis of Hoekert et al., 2007). Together with difficulties in facial emotion recognition, this deficit contributes to the impairment of emotional processing in schizophrenia, and may affect functional outcomes (Kee et al., 2003). Furthermore, evidence suggests that positive symptoms have a significant association with this deficit (Rossell and Boundy, 2005), particularly with positive emotions (Tseng et al., 2013). Similar neurofunctional underpinning for the emotional prosodic recognition deficits and the hallucinatory experiences highlights its importance for exploring the pathogenesis of this psychotic experience (Alba-Ferrara et al., 2012).

Similar to the facial emotion recognition deficits, these difficulties are also found in patients with first episode psychosis (FEP) (Edwards et al., 2001, Kucharska-Pietura et

al., 2005) and who are at clinical or familial high risk (Kee et al., 2004), with a less severe form than those seen in patients with schizophrenia or other psychotic disorders (Addington et al., 2012, Amminger et al., 2012a, Amminger et al., 2012b, Thompson et al., 2012). It seems that prosodic voice processing is also an important indicator for the subtle emotional processing difficulties in UHR subjects. Further inspection into the alteration in the neural substrate underlying emotional prosodic recognition in both FEP and UHR subjects will help to elucidate the nature of emotional processing difficulties in prodromal and early psychosis.

6.1.1 The neural correlates of emotional prosodic voice processing

Recent advances in neuroimaging techniques have allowed the delineation of neural networks underlying emotional prosody recognition. These networks are involved in the extraction of acoustic cues containing emotional information, evaluating the emotional valence, and integrating prosodic cues with other relevant information. This network of areas includes superior temporal cortex, inferior frontal cortical areas, insula, basal ganglia and to a lesser extent, the amygdala (See meta-analysis by Witteman et al., 2012).

A multi-stage model of emotional prosody perception has been proposed. In the initial phase the middle superior temporal gyrus (mSTG) is responsible for extracting basic acoustic properties from the speech signal (stage 1). This is followed by integration of

acoustic information into an emotional percept in the posterior superior temporal gyrus (pSTG, stage 2). Finally the emotional prosody is explicitly evaluated and integrated with other relevant information by bilateral inferior (IFG) and orbital frontal cortex (stage 3) (Wildgruber et al., 2009, Bruck et al., 2011). The amygdala and basal ganglia play a role in the detection of emotional salience and the evaluation of valence at this final stage (Further discussion see review of Witteman et al, 2012).

6.1.2 The neural correlates of prosodic emotional processing in schizophrenia and FEP

Few studies have focused on neural substrates for the prosodic emotional processing in schizophrenia. Patients with schizophrenia show significant deficits in recognising prosodic emotions, which correlates with lower fractional anisotropy values within primary and secondary auditory pathways, orbitofrontal cortex, corpus callosum, and peri-amygdalar white matter (Leitman et al., 2007). This suggests that structural disturbances at the level of primary auditory cortex contribute to the deficit of prosodic emotional processing in schizophrenia.

Further aberrant brain organization may underlie difficulties with prosodic emotion recognition in schizophrenia. Mitchell et al. (2004) found that patients with schizophrenia activate the left STG to a lesser extent than healthy controls, but with greater activation in left MTG, left parahippocampal gyrus and right precentral gyrus during passively listening to meaningless syllables with emotional prosody.

Furthermore, those patients who experienced auditory hallucinations displayed an atypical lateralisation, resulting in impaired emotional processing, although this pattern of the lateralisation remains controversial in the literature(Mitchell et al., 2004, Mitchell and Crow, 2005, Bach et al., 2009) and may be related to the verbal complexity (Mitchell and Crow, 2005). To summarise, though very few studies have directly focused on the neurofunctional correlates of emotional prosodic recognition in subjects with psychosis and healthy controls, existing studies suggest that altered activation in the temporal and parietal area, in particular along primary auditory pathways and secondary auditory pathways (mSTG and pSTG) may underlie the deficit of prosodic recognition in schizophrenia, and that atypical lateralisation may also play a role.

6.1.3 The neural correlates of emotional processing in UHR

As described in the previous chapter, UHR subjects show subtle alterations in brain function and structure relative to healthy controls. Alterations in regions crucial for emotional processing have also been shown (Fusar-Poli et al., 2007b, Smieskova et al., 2010), but so far no functional studies have explicitly examined prosodic recognition in an UHR cohort. Areas responsible for prosodic decoding (mSTG), integration (pSTG) and evaluation (Amygdala, parahippocampal area, IFG) in healthy controls are also areas where structural and functional alterations are reported in UHR subjects (Seiferth et al., 2008, Broome et al., 2009, Fusar-Poli et al., 2011, Mechelli et al., 2011, Tognin et al., 2013).

In addition, as described in the previous chapter, caudate nucleus is also an area of interest in terms of the processing of negative emotions (Carretie et al., 2009) and DA synthesis in UHR (Egerton et al., 2011, Howes et al., 2011)

6.1.4 Aims and Hypotheses

The aims of the study described in this chapter were to examine the neural substrate of prosodic emotion recognition in UHR and FEP groups relative to healthy control subjects using a prosodic emotion recognition task. It is hypothesised that:

6.1) FEP subjects would show deficits in prosodic emotion recognition, relative to healthy controls, especially for prosodies with negative emotions. Alterations in the neural substrate are expected in regions related to prosodic emotion recognition, particularly STG, amygdala and IFG.

6.2) UHR subjects would show intermediate deficits (i.e. performance and activation deficits relative to healthy controls that are less severe than in FEP patients) during the prosodic voice task particularly for negative emotions.

6.3) Altered activation in the striatum/caudate nucleus relative to controls during negative comparing to the positive emotions (described in hypothesis 5.3, page 94) would also be seen in UHR subjects relative to controls during the prosodic voices.

6.2 Materials and Methods

6.2.1 Participants

Participant demographics, clinical data, and estimated IQ are presented in Table 5.1 (page 106). The study obtained ethical approval as previously described in Section 5.2.1(page 95). Subject inclusion and exclusion criteria were identical to those described in chapter 5 as were the subjects who participated i.e. eighteen subjects with FEP, sixteen UHR subjects and twenty-one healthy participants.

6.2.2 Data Acquisition

Neurocognitive Assessment and Estimate of Premorbid IQ, Clinical Symptom Profiles, self-report questionnaires, and the magnetic resonance imaging procedure are identical to those described in sections 5.2.1.1 – 4.

6.2.3 Data Analysis

6.2.3.1 Demographic, neurocognitive and clinical measurement Data

The analyses for demographic, neurocognitive and clinical measurement were described in 5.2.3.1 (page 100).

6.2.3.2 Prosodic Emotional Voice Recognition: Behavioural Analysis

Based on their responses during the task, the mean proportion of accuracy and mean reaction times were calculated for each subject. As described in the previous chapter,

the participants who were entered into the analysis comprised 20 HC, 16 UHR and 18 FEP subjects, with one healthy subject excluded due to incomplete data collection.

Accuracy and RT were analysed using separate repeated measures analyses of variance (RM-ANOVAs). Only the voice clips conveying high intensity emotional prosody were entered into the analysis since the low-intensity clips served as near-neutral contrasts in the analysis of functional imaging data. The emotional category (happy, sad, fear) was entered as the within-subject variable. Diagnostic group was entered as the between subject variable.

6.2.3.3 Functional MRI

The preprocessing procedure and the first level analysis were identical to that described for the dynamic faces experiment detailed in 5.2.3.3 (page 102).

Using the GLM parameter estimates obtained for all brain voxels, five contrasts of interest: Overall emotions in Voice, Happy-near Neutral Voice, Sad-near Neutral Voice, Fearful-near Neutral Voice, and Positive – Negative voices were computed from the six voice conditions out of ten experimental conditions described in Chapter 5: 5) High-intensity Happy Voice 6) Low-intensity Happy Voice 7) High-intensity Sad Voice 8) Low-intensity Sad Voice 9) High-intensity Fearful Voice 10) Low-intensity Fearful Voice. The low-intensity Voice conditions served as near-neutral baselines for each high intensity voice contrast. Second level analysis of the contrast images was

performed using an independent ANCOVA F-test implemented in SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) running under Matlab 7.1 (Math Works, Natick, MA, USA), with age entered as a covariate and statistical inferences made at $p<0.05$ after FWE correction for multiple comparisons at cluster level.

In order to correct for multiple comparisons in interpreting these results, a bilateral STG ROI activated in healthy controls during the prosodic emotion recognition, was used. The coordinates were chosen from the meta-analysis by Witteman et al. (2012). Considering that neurofunctional alteration in schizophrenia patients during emotional processing may exhibit similar changes across modality, areas showing significant group differences for facial emotional recognition where also examined (i.e. fusiform area, left amygdala, and right side lentiform nucleus). Those coordinates used in the facial expression study, i.e. bilateral fusiform gyrus, left amygdale and right lentiform gyrus (Li et al., 2010) were also included to test for between group differences during emotional processing. Additionally, the caudate nucleus (Carretié et al., 2009) was also chosen as a region of interest for testing hypothesis 6.3 - relating to valence-specific interaction effects during the negative emotional voices relative to positive emotional voices.

A small volume correction (SVC) with a sphere of 10 mm radius was used according to the coordinates of previous studies, including left STG (-62, -22,1), right STG (49, -23, 6), left fusiform gyrus (-39, -65, -13), Right fusiform gyrus (38,-64,-10), left amygdala (-21, -7, -8), right lentiform gyrus (22, -3, -5), left caudate body (-18, -2, 24) and right

caudate body (16, 4, 18). All the coordinates reported were described using a standard Montreal Neurologic Institute (MNI) coordinate system.

6.2.3.4 Correlational Analysis

Pearson's correlation analyses were performed to examine the relationship between the peak level of activation within ROIs showing group differences during prosodic voices task and behavioural performance (accuracy and reaction time), clinical symptoms (including total scores, positive symptom scores and negative symptom scores for PANSS in both UHR and FEP groups and CAARMS severity scale in the UHR group).

6.3 Results

6.3.1 Prosodic voice recognition performance

Accuracy: A repeated measures ANOVA showed a significant main effect of emotion (F =17.84, df =2, p <0.001). The main effect of emotion showed that accuracy was greatest for sad, lowest for fear and intermediate for happy. The main effect of diagnostic group was significant (F =12.89, df =2, p<0.001), mainly driven by better accuracy in NC than in UHR and in FEP. The diagnostic group*emotion interaction was non-significant (F =0.90, df =4, p=0.47). The accuracy for individual emotions are reported in detail in Figure 6.1(a).

Reaction Times: A repeated measures ANOVA showed a significant main effect of emotion (F =47.84, df =2, p <0.001), with the reaction time shortest for happy, longest for fear and intermediate for sad. The main effect of diagnostic group (F =0.69, df =2, p =0.51), and the diagnostic group*emotion interaction (F =1.06, df =4, p =0.38) were non-significant. The reaction times for individual emotion are reported in detail in Figure 6.1(b).

Figure 6.1 (a)

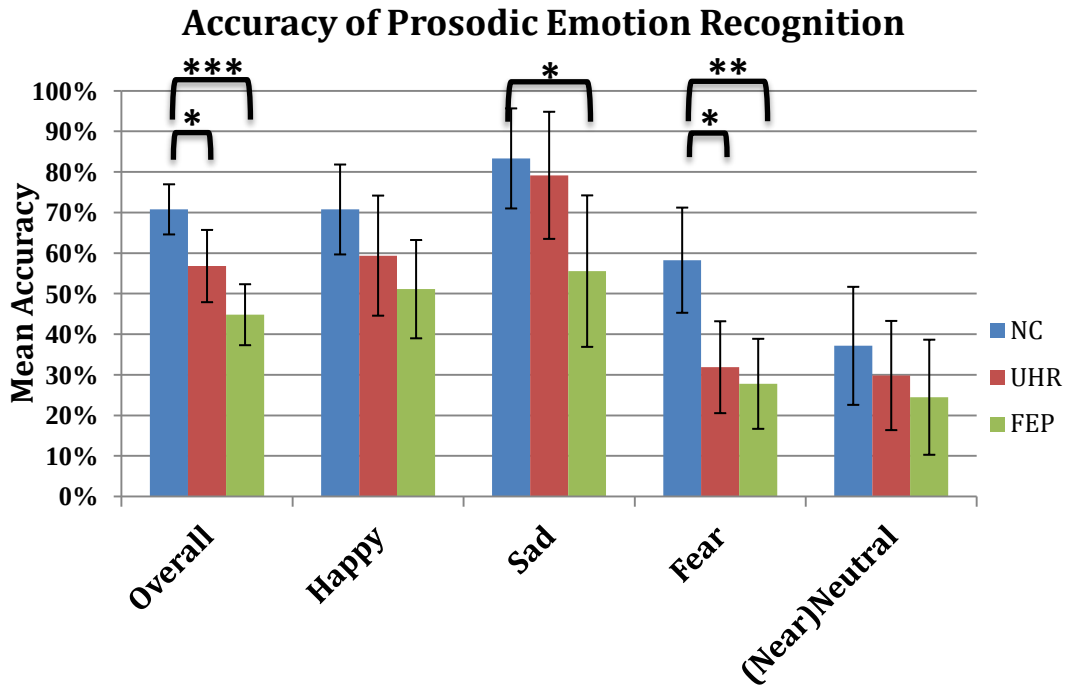


Figure 6.1 (b)

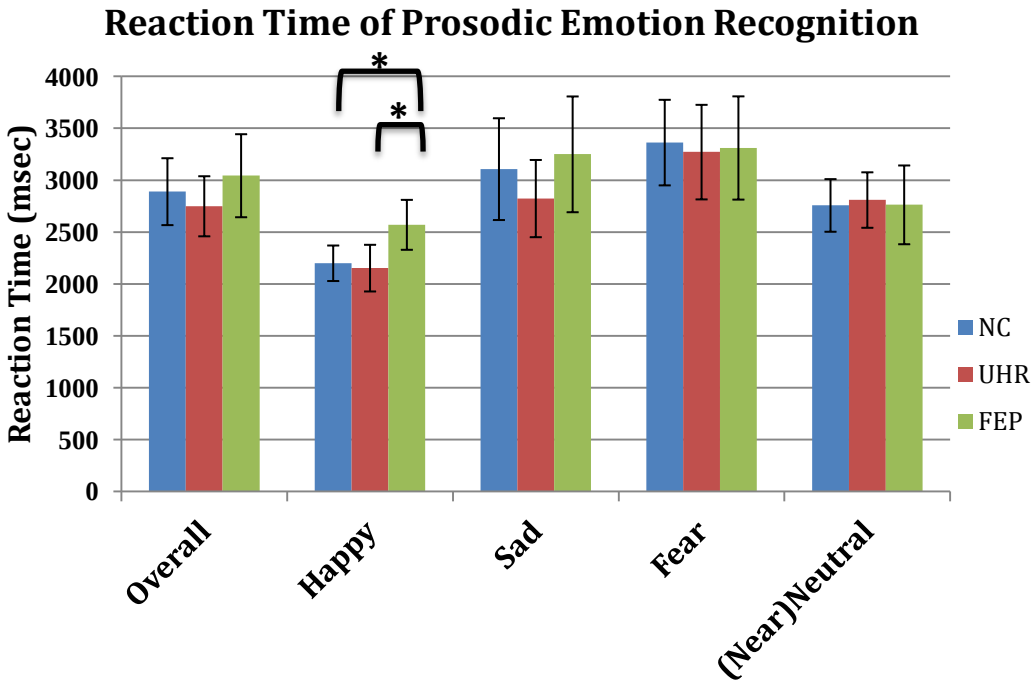


Figure 6.1 Graph showing (a) mean accuracy, and (b) mean reaction time for each group by emotional category. Near-neutral category shows the average performance across three emotions for low-intensity emotional stimuli.

6.3.2 Functional MRI

6.3.2.1 MAIN EFFECT OF TASK

Emotional voice > Near-neutral voices

Whole brain analysis showed a main effect of emotion (happy, sad, fear) in the right supramarginal gyrus (62, -48, 32), left hippocampus (-26, -14, -18), and left precuneus (0, -76, 46) ($p<0.001$, uncorrected; See Figure 6.2).

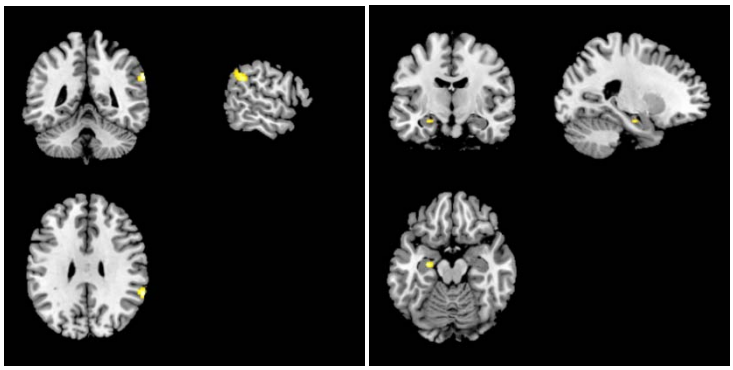


Figure 6.2 Peak BOLD activation for prosodic emotion task effect across groups.

Positive versus Negative voices

Whole brain analysis showed an effect of positive > negative emotion in the anterior cingulate (4, 16, 26) ($p<0.001$, uncorrected, see Figure 6.3). The negative > positive emotion across groups was non-significant.

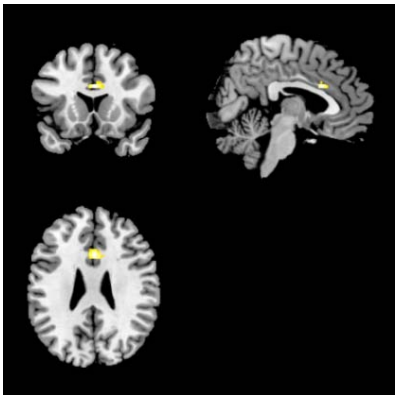


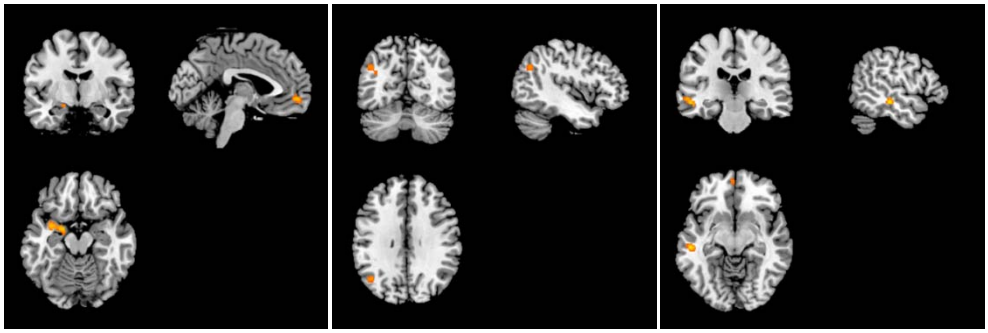
Figure 6.3 Peak BOLD activation for positive > negative emotion effect across groups

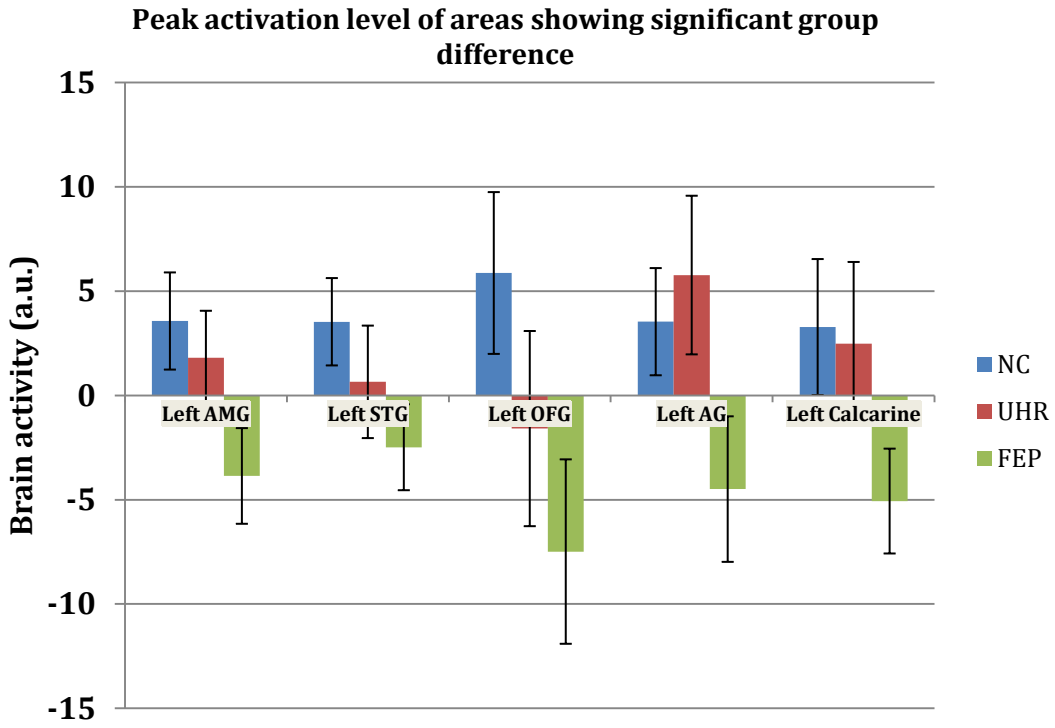
6.3.2.2 MAIN EFFECT OF GROUP

Using whole brain analysis, an effect of diagnostic group across emotions was seen in the left amygdala extended to the left insular area, left superior temporal gyrus, medial orbital frontal lobe, left angular gyrus, and left calcarine ($p<0.001$, uncorrected, cluster size >50, see Table 6.1 and Figure 6.4).

Activated brain regions	No. of voxels				Maximum F values	*Z*	*p* value	
		x	y	z			uncorrected	FWE corrected
Left amygdala	129	-32	0	-18	15.24	4.33	0.00001	0.131
		-20	-2	-14	12.69	3.96	0.00004	0.413
Left superior temporal gyrus	92	-52	-26	-12	13.00	4.01	0.00003	0.364
Medial orbital frontal gyrus	100	-2	54	-6	11.54	3.77	0.00008	0.627
Left middle occipital gyrus	52	-32	-84	26	11.08	3.69	0.00011	0.718
		-22	-82	18	8.06	3.10	0.00096	0.999
Left lingual gyrus	62	-6	-80	8	10.22	3.54	0.00020	0.869
		-14	-78	8	8.99	3.30	0.00048	0.982
Left angular gyrus	59	-44	-64	32	10.19	3.53	0.00020	0.873
		-36	-64	26	9.12	3.33	0.00044	0.976

Table 6.1 Coordinates showing group main effects ($p<0.001$, uncorrected)





*AMG = Amygdala; STG = Superior Temporal Gyrus; OFG = Orbital Frontal gyrus; AG = Angular Gyrus

Figure 6.4 Peak BOLD activation for group effect across emotions.

The main effect of group was mainly driven by decreased activation in the FEP patients relative to HC in the left amygdala extending to left insula, left middle occipital lobe extending into calcarine area, and cerebellum. ($p<0.05$, FWE corrected; see Figure 6.5)

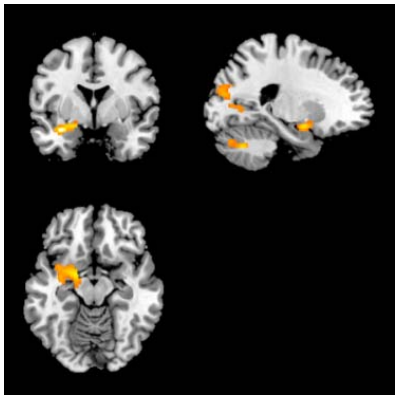


Figure 6.5 Peak BOLD activation for group effect across emotions, NC>FEP.

6.3.2.3 GROUP × TASK INTERACTION

Whole brain analysis showed a group-by-emotion interactions between the NC and the UHR group in caudate body (p<0.001, uncorrected, see Table 6.2), driven by higher activation during negative emotions relative to positive emotions in NC, but the opposite pattern in UHR subjects (See Figure 6.6).

Activated brain regions	No. of voxels	x	y	z	Maximum *t* values	*Z*	*p* value uncorrected	*p* value FWE corrected
Left caudate body	16	-16	10	24	3.80	3.54	0.00020	0.80

Table 6.2 Coordinates showing group × task interactions (p<0.001, uncorrected)

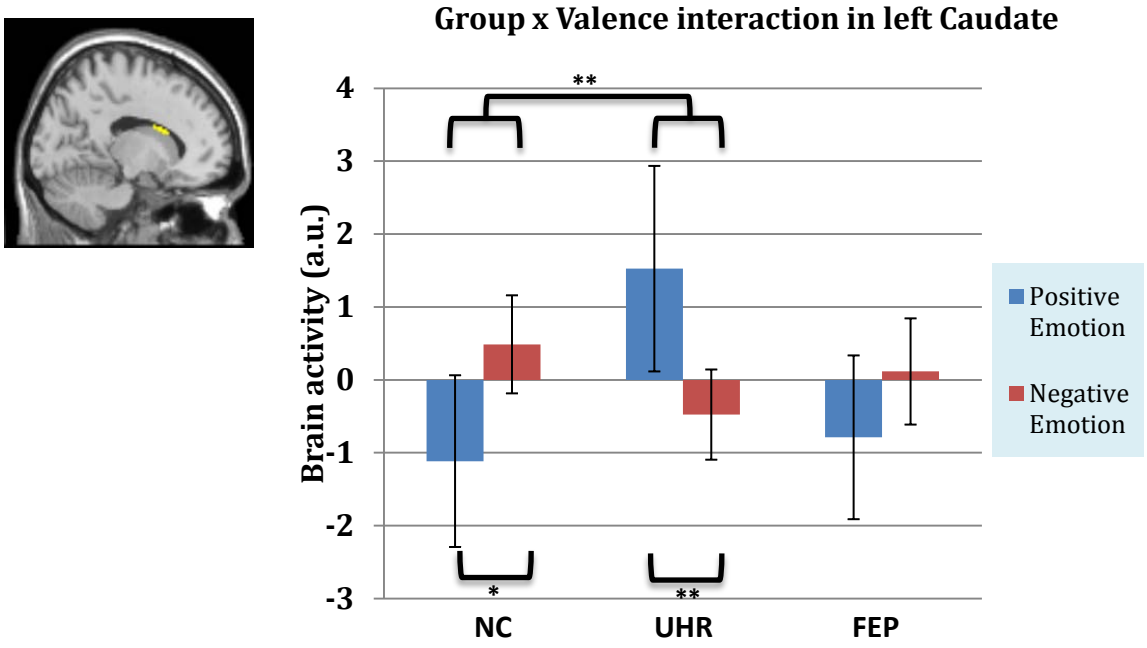
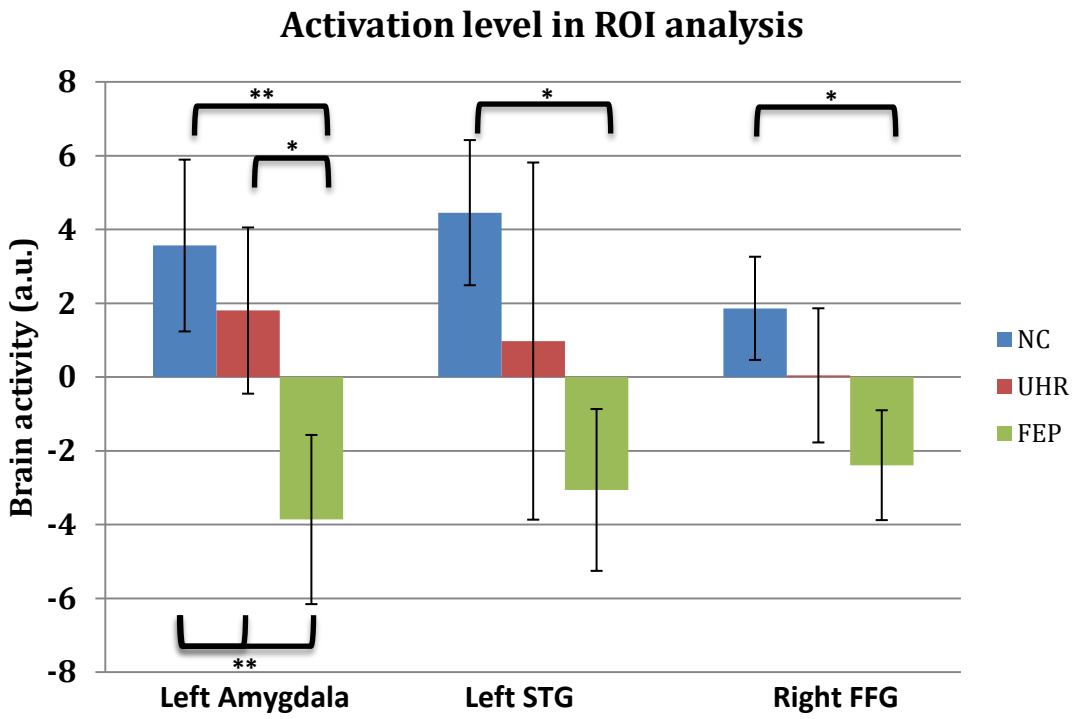


Figure 6.6 Graph showing peak BOLD activation level in the left caudate body for positive>negative effect, UHR>NC (the graph used threshold at $p<0.05$ uncorrected for better illustration).

6.3.2.4 REGION OF INTEREST ANALYSIS

Using the ROIs and coordinates described above, a trend of average effect of tasks (relative to near-neutral stimuli) across group was seen in the left amygdala (t=2.86, p=0.08). A main effect of group was seen in the left amygdala (F[2,49]=12.01, p=0.005). FEP patients had significantly lower activation in the left amygdala relative to NCs (t=5.05 p=0.003) and UHR subjects (p=0.03, t=4.20; see Figure 6.7). Higher activation levels in the left amygdala was associated with greater accuracy for fear prosody across all subjects (r=0.37, p=0.007).

In addition, FEP patients had significantly lower activation than the NC group in left STG (p=0.02, t=4.44) and right FFG (p=0.03, t=4.19; See Figure 6.7).



*STG = Superior Temporal Gyrus; FFG = Fusiform Gyrus

Figure 6.7 Graph showing peak BOLD activation level in the left amygdala, left STG, and right FFG for group effect across emotions

A significant task by group interaction between NC and UHR was seen in the left caudate area (t=3.27, p=0.03; see Figure 6.6), showing greater activation for negative relative to positive emotions in HC, but the opposite pattern in UHR subjects. Lower positive-negative activation difference (i.e. lower activation for positive emotions or higher activation for negative emotions) in the UHR group in left caudate was also associated with higher accuracy for fearful voices (r=-0.58, p=0.02). The association pattern was non-significant in the FEP group. All other interactions were non-significant. All the results reported in this section were corrected for multiple comparisons using Bonferroni correction.

6.3.3 Correlation Analysis

Using the peak eigenvalue extracted from the region of interests that showed significant effects between groups (i.e. left STG, left amygdala, and right fusiform gyrus at the current analysis), correlation analyses with clinical symptoms were performed.

Higher activation in right FFG was associated with higher positive symptom scores of CAARMS in UHR (r=0.60, p=0.02). Lower positive-negative activation difference (i.e. lower activation for positive emotions or higher activation for negative emotions) in left caudate is associated with higher PANSS total score across UHR and FEP (r=-0.42, p=0.02), mainly driven by association with negative symptoms in the FEP group (r=-0.51, p=0.04).

6.4 Discussion

The current study aimed to investigate whether emotional prosodic recognition and associated neural substrate were altered in UHR and FEP groups relative to a HC group.

6.4.1 Behavioural Results

In line with hypothesis 6.1, FEP patients were less accurate than HC group in recognising prosodic emotional voices, particularly during fear and sad emotion trials. As expected in hypothesis 6.2, the performance of UHR group was numerically intermediate between HC and FEP groups and significantly less accurate than HC,

particularly during fear emotion trials. However, the performance difference between UHR and FEP groups was not statistically significant. The overall reaction times were comparable in all three groups, though FEP patients were significantly slower when identifying happy prosodic voices. The finding of an intermediate task performance in UHR subjects, relative to HC and FEP patients, during the prosodic voices task is consistent with findings with a previous study of emotional cognition in an UHR cohort (Thompson et al., 2012) and the results of the dynamic facial recognition experiment described in the previous chapter. Furthermore, a deficit in prosodic voice recognition appear more robust than deficits in emotional face recognition, which is in line with the results reported by Amminger et al. (2012a) and meta-analyses of emotion recognition in schizophrenia (Hoekert et al., 2007, Kohler et al., 2010).

6.4.2 Whole brain analysis

A main task effect of prosodic emotional voices across groups was seen in the right supramarginal gyrus and left hippocampus. Task related activation in these regions were slightly different from previously reported neural substrates for prosodic emotional processing, which involves superior temporal cortex, inferior frontal cortical areas, insula, basal ganglia and the amygdala (Witteman et al., 2012). However, when the main task effect was restricted to healthy controls a more extensive network was revealed including bilateral posterior STG, left MTG ($p<0.05$, FWE corrected), bilateral thalamic area, amygdala and medial orbital frontal area ($p<0.001$, uncorrected). It is likely that when the main effect of task is specified across all subjects in the model, that

the increased variance introduced by the UHR and FEP groups decrease the robustness of the effect in some areas.

Broadly consistent with hypothesis 6.1, a significant group effect was observed during the prosodic voice task particularly for negative emotions. A group difference was seen in a left sided network of regions including left amygdala, left insular area, left thalamus, left middle temporal gyrus, medial orbital frontal lobe, left calcarine and cerebellum. The group effect in these regions was mainly driven by the reduced activation the FEP relative to HC group.

Although there were no significant activation differences in the inferior frontal areas across groups, there were clusters showing decreased activation in FEP patients relative to NC in left IFG (-30, 26, -14, cluster size=16) and right insula extended to right IFG (32, 20, -14, cluster size=127), which did not survive correction for multiple comparisons. This suggests a trend for reduced activation in the bilateral inferior frontal and right insular area.

6.4.3 ROI Results

A trend of average effect of tasks across group was seen in the left amygdala. Activation in the left amygdala was positively associated with the accuracy of fearful prosody across groups, suggesting that the amygdala is responsible for processing

fearful emotions (Adolphs et al., 1995, Adolphs, 2008) and that an increased activation level in this region facilitates accurate prosodic recognition.

As expected in hypothesis 6.1, the ROI analysis revealed a significant decrease in activation in FEPs compared to HCs in the left STG and left amygdala, areas known to be involved in emotional processing of prosodic voice (Leitman et al., 2007, Witteman et al., 2012). The differences between NC and FEPs remain significant even when the accuracy was entered as a covariate (F=11.83, p=0.002), confirming that the amygdala hypoactivation was not secondary to poor recognition accuracy in FEP patients.

Activation in the UHR group was broadly intermediate between NC and FEP; although the activation in the UHR was significantly higher than in FEP patients in the left amygdala it was not significantly reduced relative to HCs. This finding is in line with those reported by Rasetti et al. (2009) who found that patient with schizophrenia demonstrated reduced amygdala reactivity to negative emotional stimuli whilst unaffected siblings showed a pattern that was not significantly different from that of healthy comparison subjects. The authors suggest that rather than a genetic trait, hypoactivation of the amygdala in processing fearful representations is more likely related to the disease state (Rasetti et al., 2009).

A group by task interaction in the caudate body revealed a pattern of higher activation for positive relative to negative emotions in UHR subjects, whilst the opposite pattern

was seen in the HC and the FEP groups. The caudate body is known to be sensitive to negative emotional stimuli (Carretie et al., 2009), but for UHR subjects this region was more active during positive relative to negative prosodic voice trials. Interestingly, fearful prosody recognition was reduced in UHR subjects relative to controls. It is tempting to speculate that reduced fear prosody recognition accuracy in UHR subjects is due to a positive-negative emotion confusion brought about by aberrant caudate body activation. The observation that UHR subjects with greater activation difference for positive relative to negative emotions in caudate body were less accurate for fearful voices further supports this speculation.

Activation in right FFA was also significantly lower in FEP than in NC. FFA is responsible for facial emotion recognition, although studies also report significant activation of this area during prosodic recognition (Alba-Ferrara et al., 2012). The FFA may be involved in the early perceptual processing of emotional information, and dysfunction of FFA may be a primary deficit responsible for emotional processing problems in the psychosis.

6.4.4 General discussions

Overall, consistent with the hypothesis, the findings of the current study suggest that FEP patients were significantly impaired relative to HC in prosodic emotion recognition. Analyses of fMRI data during emotional prosodic voice recognition showed a reduced activation in a left sided temporal-parietal-limbic circuit in FEPs. The UHR subjects

also demonstrate overall intermediate recognition accuracy between HC and FEP groups, and significantly reduced fear prosodic recognition relative to healthy controls. The neural substrate of prosodic recognition was largely unaffected, but the activation was intermediate between HC and FEP in areas related to emotional processing, including STG, FFG and amygdala. Furthermore, altered activation patterns in the right fusiform gyrus and caudate nucleus were seen in UHR subjects relative to HC. These results suggest that subtle neurofunctional alterations may already exist in the UHR, and might be related to the misattribution of the emotions.

Deficits in emotional recognition of prosodic stimuli was more predictive of positive symptom dimensions in schizophrenia, and can potentially be a neurocognitive measurement for objectively evaluating the symptom severity of auditory hallucination, as similar brain region are involved, in particular superior temporal area (Alba-Ferrara et al., 2012, Tseng et al., 2013). A hypothetical model purposed by Alba-Ferarra et al. suggests that dysfunctional superior temporal gyrus and disrupted functional connectivity with anterior cingulate cortex may underlie the aberrant integration of prosodic features and misidentification of the source of the voice, which consequently contribute to hallucination formation. The result that FEP patients had significantly lower activation in left STG partially supported this hypothetical model. Further connectivity analyses between STG, anterior cingulate gyrus and amygdala are warranted to confirm the model.

6.5 Limitations

A major limitation for the prosodic voices task was the lack of an emotionally neutral voice condition. Because there were no neutral prosodic stimuli in the original design of DANVA2, introducing new stimuli outside the original dataset may have unwanted confounding effects. However, the regions activated by the task in HC were similar to those described in the prosodic processing literature (Witteman et al., 2012).

6.6 Conclusion

To conclude, consistent with the majority of the literature, the findings of the current study suggest that FEP patients were significantly impaired relative to HC in prosodic emotion recognition. We have shown that UHR subjects also demonstrate reduced prosodic recognition relative to healthy controls.

Analyses of fMRI data during emotional prosodic voice recognition showed a reduced activation in a left sided temporal-parietal-limbic circuit in FEPs that was associated with negative symptoms and lower social functioning.

In UHRs, despite significantly decreased accuracy, the neural substrate of prosodic recognition was largely unaffected. However, altered activation patterns in the right fusiform gyrus and caudate nucleus were seen in UHR subjects relative to HC. These alterations may be responsible for emotion misrecognition, which might be related to the misattribution of the emotions and the existence of positive symptoms in the UHR.

Chapter 7

Audiovisual integration of Emotional Stimuli in Early and Prodromal Phase of Psychosis

7.1 Introduction

Multisensory integration (MSI) is a crucial complex perceptual process to effectively obtain important and holistic information in everyday life. The available literature reviewed in Chapter 2 indicates that patients with schizophrenia demonstrate impairments in the integration of non-emotional audiovisual stimuli, especially of complex linguistic stimuli (de Gelder et al., 2003, See Chapter 2), as well as altered integration of emotional stimuli. Although the underlying mechanisms of these deficits remain unclear, attentional deficits in schizophrenia patients have been considered as a possible overarching factor (de Jong et al., 2010). Overall, impaired MSI is likely to be associated with the pronounced emotional processing and social cognition deficits seen in schizophrenia patients.

Although the superior temporal sulcus (STS) and the adjacent area has been considered crucial during both multisensory non-emotional and emotional integration (Ethofer et al., 2006a, see below 7.1.1), few neuroimaging studies have focused on MSI deficits in schizophrenia (Szycik et al., 2009, Muller et al., 2013b, Straube et al., 2013, Szycik et al., 2013). Moreover, research into emotional MSI in prodromal and/or early phases of psychosis is sparse. Given that MSI may underlie the early presentation of the

emotional and interpersonal disturbance in the trajectory of development of psychosis, further research into MSI in the early phase of psychosis is warranted.

7.1.1 The neural correlates of multisensory integration of non-emotional and emotional stimuli

A congruent facilitation effect (see Chapter 2) has been demonstrated while healthy subjects perceive emotionally congruent information from a different sensory modality (de Gelder and Vroomen, 2000). Cortical areas activated during non-emotional multisensory integration are also activated by bimodal emotional stimuli, presumably as part of a general integration process. The superior temporal gyrus (STG) and STS in particular are active during bimodal integration conditions (Ethofer et al., 2006b).

Kreifelts et al. (2009) further elucidate functional differentiation between sub-regions of STS during an emotional audiovisual integration task. The authors report that the trunk section of STS is more sensitive to emotional auditory stimuli, the posterior terminal ascending branch of posterior STS is more sensitive to emotional facial stimuli, and increased activation during audiovisual integration is observed in the bifurcation of posterior STS, situated between the auditory- and visual-sensitive areas (See Figure 7.1, page 154).

The amygdala also plays a role in modulating the cross-modal integration and has been shown to be activated when subjects are presented with congruent fearful face-voice pairs stimuli (Dolan et al., 2001). Increased activation is also seen in inferior frontal gyrus, parahippocampal gyrus and amygdala during audiovisual emotional integration relative to single modality stimuli alone (Park et al., 2010), and the posterior insula is activated to a greater extent by congruent audiovisual information relative to incongruent information (Pourtois et al., 2005, Ethofer et al., 2006b). Furthermore, enhanced connectivity between the left amygdala and right fusiform cortex is seen during emotional multisensory integration (Ethofer et al., 2006b).

Positron Emission Tomography (PET) and fMRI studies have shown that multi-sensory integration of different emotions is mediated by distinct neuroanatomical substrates (Pourtois et al., 2005, Park et al., 2010). For example, Park and colleagues (Park et al., 2010) report anger-specific activation in the posterior cingulate, fusiform gyrus, and cerebellum, while happiness-specific activation is seen in the medial temporal gyrus (MTG), hippocampus, claustrum, inferior parietal lobule, cuneus, middle frontal gyrus, and anterior cingulate during bimodal integration. This implies each emotion category may rely on separate networks to integrate bimodal information. In sum, cortical areas surrounding the STS are involved in both non-emotional and emotional integration of audiovisual stimuli, while posterior insula and the amygdala show additive responses when emotions are expressed in a congruent way, especially for fearful stimuli (Ethofer et al., 2006a).

Multisensory emotional conflict is usually studied by presenting incongruent emotional cues in different sensory modalities such as a sad facial expression with an angry voice (de Gelder et al., 1999). Similar to non-emotional conflict, emotional conflict has been linked to activation in the anterior cingulate cortex (ACC) both within (Haas et al., 2006, Ochsner et al., 2009, Wittfoth et al., 2010) and between sensory modalities (Muller et al., 2011). Incongruent emotional valence between faces (happy, fear) and sounds (laughing, screaming) has also been shown to increase activation in the middle cingulate cortex, right superior frontal cortex, right supplementary motor area and right temporoparietal junction (Muller et al., 2011). The amygdala, which is also involved in the integration process of emotional information from different sensory modalities (Dolan et al., 2001), does not show significant activation during incongruent emotional stimuli. These results indicate that in healthy volunteers, incongruent emotional stimuli activate a cingulate-fronto-parietal network involved in conflict monitoring and resolution (Muller et al., 2011) with less activation in emotion-related networks.

7.1.2 The neural correlates of multisensory integration in FEP and UHR

The neural substrate of multisensory processing of both emotional and non-emotional stimuli in schizophrenia has been systemically reviewed in Chapter 2. To date, no study has examined the neural substrate of multisensory processing in subjects with first episode or ultra high risk (UHR) states for psychosis.

Similar to patients with schizophrenia (Gur et al., 2002a, Hempel et al., 2003, Williams et al., 2004), patients with early psychosis and UHR subjects show anatomical and functional abnormalities in the neural circuitry associated with emotional processing relative to healthy controls. In particular, altered activation in the right lingual and fusiform gyrus and altered prefronto-limbic functional connectivity have been reported (Seiferth et al., 2008, Modinos et al., 2010, Smieskova et al., 2010).

7.1.3 Aims and Hypotheses

The aims of the study described in this chapter were to examine the neural substrate of multisensory emotional integration in UHR and first episode psychosis (FEP) groups relative to healthy control subjects using a multisensory emotional integration task. It was hypothesised that:

7.1) During a multisensory emotional recognition task, (a) FEP subjects would show reduced recognition accuracy during emotionally congruent trials. (b) FEP patients will also show less facilitation effect from emotionally congruent trials relative to emotionally incongruent trials (accuracy of emotionally congruent trials – accuracy of emotionally incongruent trials) relative to healthy controls.

7.2) UHR subjects will demonstrate intermediate performance between controls and FEP in recognition accuracy, facilitation and interference effects.

7.3) During a functional Magnetic Resonance Imaging (fMRI) study using the multisensory emotional recognition task, FEP subjects would show altered activation in regions related to integration of audiovisual emotional information, particularly

bifurcation section of pSTS (anterior section of pSTS, apSTS) and amygdala during both congruent and incongruent emotional stimuli relative to healthy controls. During incongruent trials, activation in areas related to conflict detection and conflict resolution would also be altered in FEP subjects.

7.4) UHR subjects would demonstrate intermediate alteration in the magnitude of activation between controls and FEP patients.

7.2 Materials and Methods

7.2.1 Participants

The study obtained ethical approval as previously described. The inclusion and exclusion criteria were identical to those described in Section 5.2.1 (page 95).

The Participants have been described in details in Chapter 5. Briefly, the participants comprised of eighteen FEP patients, sixteen UHR subjects and twenty-one healthy controls. Participants' demographics, clinical data, and estimated IQ are presented in Table 5.1 (page 106).

7.2.2 Data Acquisition

Neurocognitive assessment and estimated IQ, clinical symptom profiles, and self-report questionnaires are the same as those described previously in sections 5.2.2.1 – 3 (page 97 – 98).

7.2.2.1 Magnetic Resonance Imaging

Neuroimaging was conducted using a 1.5T MRI scanner (Sigma, LX-GE, Milwaukee, USA) at the Maudsley Hospital, London. The emotional recognition paradigm for fMRI used in the current experiment was a single modality emotion recognition task.

There were 96 audiovisual multisensory trials comprising 24 congruent trials (3 conditions: Happy-Happy, Sad-Sad, Fear-Fear), 24 neutral-emotional pairs (3 conditions: Neutral-Happy, Neutral-Sad, Neutral-Fear), and 48 incongruent pairs (6 conditions: Happy-Sad, Happy-Fear, Sad-Happy, Sad-Fear, Fear-Happy, Fear-Sad). The mean length of stimuli was 4.2 seconds. A one-second inter-stimuli interval with a fixation cross in the centre of the screen followed each stimulus. All multisensory trials were pseudo-randomly arranged in a session. During the task, subjects' responses (accuracy and reaction time) were recorded for subsequent performance analysis.

Functional images were acquired using a TR of 3000 ms, a TE of 40 ms, a flip angle of 90°, a slice thickness of 2.5mm with 0.5mm gap, a field of view of 24cm^2 and a 64x64 matrix. In total, 46 axial slices parallel to the anterior commissure–posterior commissure (AC-PC) line were collected for each subject. In total, 232 image volumes were generated in the single modality task for each participant.

7.2.3 Data Analysis

7.2.3.1 Demographic, neurocognitive and clinical measurement Data

The analyses for demographic, neurocognitive and clinical measurement were described in 5.2.3.1 (page 100).

7.2.3.2 Audiovisual emotional Recognition and integration: Behavioural Analysis

Based on their responses during the task, the mean proportion of accurate response and mean reaction times were calculated for each subject. One UHR subject was excluded due to incomplete data collection, resulting in an UHR group of seventeen subjects.

The emotional recognition paradigm for fMRI used in the current experiment was a dual- modality emotion recognition task established in the pilot and behavioural studies (See Chapter 3 and 4). The paradigm is adapted to demonstrate multisensory integration process by contrasting congruent trials (e.g. happy face and happy voice) to incongruent trials (e.g. happy face and sad voice). Participants were instructed to base their judgment on the voice and choose between four emotional categories (happy, sad, fear,

and neutral) via a button box as quickly as possible before the clips ended. Accuracy and reaction times (RT in milliseconds) were recorded.

Accuracy and RT were analysed using separate repeated measures analyses of variance (RM-ANOVAs). The congruency (congruent, incongruent, emotional-neutral pairing) was entered as the within-subject variable. Diagnostic group was entered as a between subject variable.

7.2.3.3 Functional MRI

The pre-processing procedures and the first level analyses were identical to the Facial experiment and have been described in 5.2.3.3 (Page 102). The analyses were limited to the dual modality task only and compared the accuracy and RTs of congruent, emotional-neutral, and incongruent pairs. Performance of single modality emotional recognition task was not included in the analyses at the current stage.

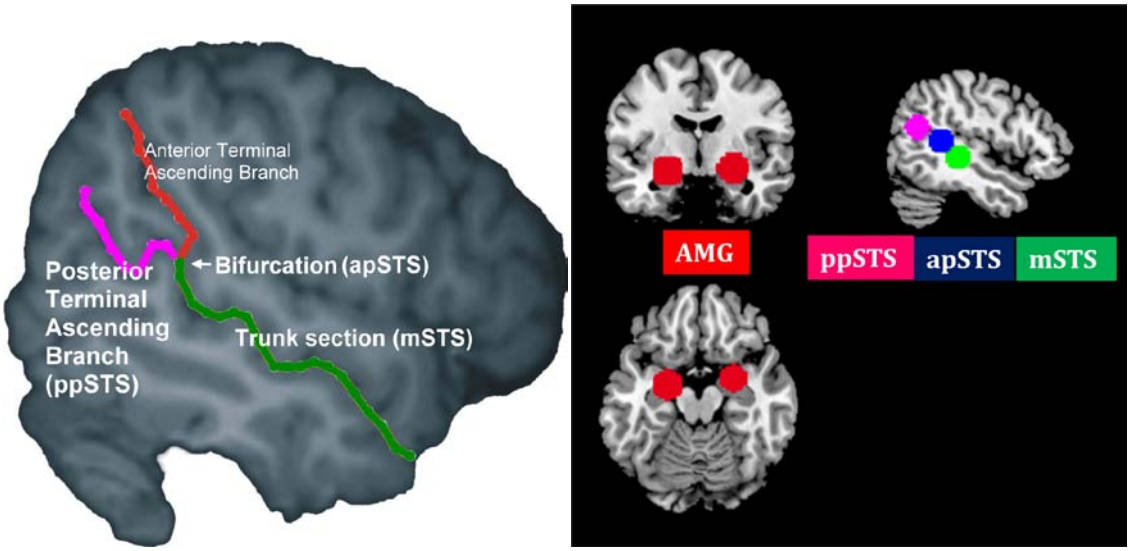
Using the GLM parameter estimates obtained for all brain voxels, twelve multisensory experimental conditions were modeled in the first level analyses: 1) Happy Voice – Happy Face, 2) Sad Voice – Sad Face, 3) Fear Voice – Fear Face, 4) Happy Voice – Neural Face, 5) Sad Voice – Neutral Face, 6) Fear Voice – Neutral Face, 7) Happy Voice – Sad Face, 8) Happy Voice – Fear Face, 9) Sad Voice – Happy Face, 10) Sad Voice – Fear Face, 11) Fear Voice – Happy Face, 12) Fear Voice – Sad Face. Two contrasts of interest, congruent emotional trials versus baseline (from condition 1-3) and

incongruent emotional trials versus baseline (from condition 7-12), were computed from these experimental conditions. The fixation cross served as baselines for each contrast. Due to the unequal number of regressors in each condition, each regressor was weighted by 1 divided by regressor numbers in the contrast (i.e. 1/3 for each regressor in the Congruent contrast and 1/6 for each regressor in the Incongruent contrast). Second level analyses of the contrast images were performed using a Group (NC, UHR, FEP; independent between levels) by Congruency (Congruent and incongruent; non-independent between levels) ANCOVA in SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) running under Matlab 7.1 (Math Works, Natick, MA, USA), with age entered as a covariate and statistical inferences made at $p<0.05$ after FWE correction for multiple comparisons at cluster level. The group by task interaction was examined in three levels separately: interactions between NC and FEP by task; interactions between NC and UHR by task, and interactions between UHR and FEP by task.

The bilateral anterior parts of posterior STS (apSTS), which were activated during audiovisual emotional stimuli in healthy controls, were selected as regions of interest. Coordinates used were based on a study in healthy volunteers by Kreifelts et al. (2009). Other sub-regions of STS reported in the Kreifelts et al. study to be activated during the emotional audiovisual integration, i.e., middle STS (mSTS) and posterior terminal ascending branch of pSTS (ppSTS), were also examined.

In addition, ROIs in limbic regions reported to be activated during single modality emotional tasks (from Li et al, 2010; described in Chapters 5 and 6, see page 104) were selected to examine the brain activation for the emotional component of the task.

A small volume correction (SVC) with a sphere of 10 mm radius was used according to the coordinates of previous studies, including left mSTS (-54, -33, 1), right mSTS (54, -30, 2), left apSTS (-54, -51, 10), right apSTS (54, -42, 10), left ppSTS (-54, -63, 16), right ppSTS (54, -59, 20), left amygdala (-22, -7, -12) and right lentiform gyrus (25, -2, -10). All the coordinates reported were described using a standard Montreal Neurologic Institute (MNI) coordinate system (See Figure 7.1).



*Modified from Kreifelts et al., 2009. AMG = Amygdala; mSTS = mid-section of the STS; apSTS = anterior section of posterior STS (near bifurcation); ppSTS = posterior section of posterior STS

Figure 7.1 Graph showing sub-regions of superior temporal sulcus (STS) and selected ROIs

7.2.3.4 Correlational Analysis

Pearson's correlation analyses were performed to examine the relationship between the peak level of activation showing significant group effects in the selected ROI and behavioural performance (accuracy and reaction time), clinical symptoms (including total scores, positive symptom scores and negative symptom scores for PANSS in both UHR and FEP groups and CAARMS severity scale in the UHR group).

7.3 Results

7.3.1 Demographics and psychopathology

Demographics, neurocognitive performances and clinical measurement have been reported in Chapter 5 (page 106).

7.3.2 MERIT integration performance

Accuracy: A RM-ANOVA showed a significant main effect of congruency ($F(1,51)=79.53$, $p<0.001$), with the accuracy highest for congruent trials across all three groups (See Figure 7.2(a)). The main effect of group was significant ($F(2,51)=7.54$, $p=0.001$), mainly driven by greater accuracy in NC relative to FEP patients. The diagnostic group × emotion interaction was non-significant ($F(2,51)=0.30$, $p=0.74$). The accuracy for individual congruency statistics are reported in detail in Figure 7.2(a).

Reaction Times: A RM-ANOVA showed a significant main effect of congruency ($F(1,51)=5.16$, $p=0.027$), with the reaction time shortest for congruent trials (See Figure 7.2(b)). The main effect of group ($F(2,51)=0.65$, $p=0.63$) and the group*emotion interaction were ($F(2,51)=0.97$, $p=0.39$) both non-significant. The reaction times for individual emotion are reported in detail in Figure 7.2(b).

Figure 7.2 (a)

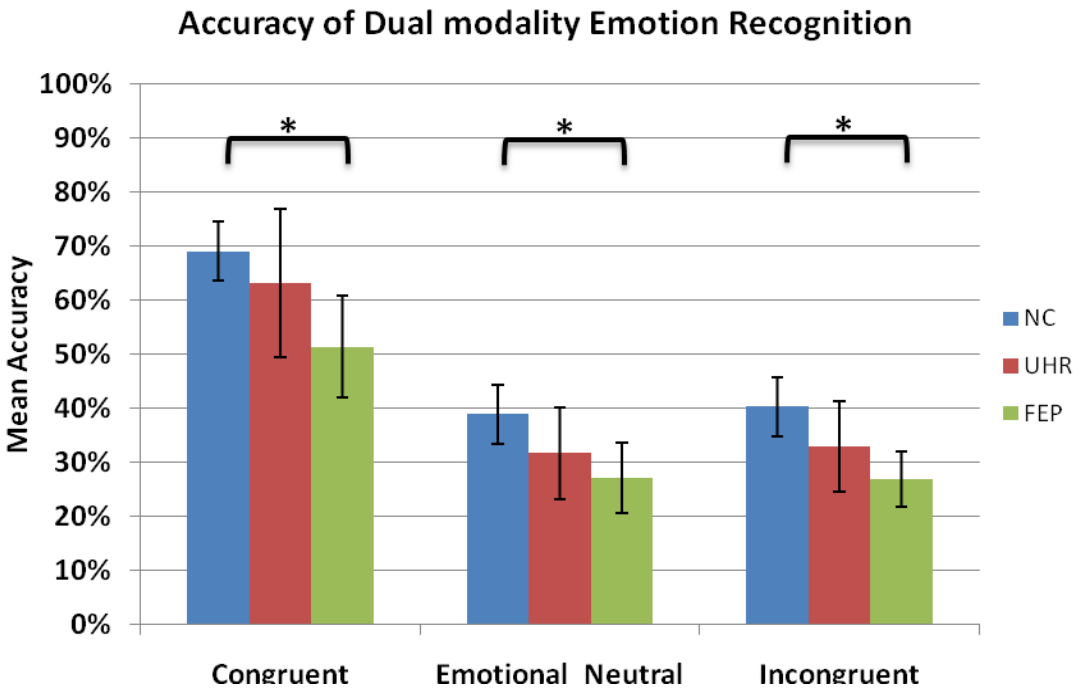
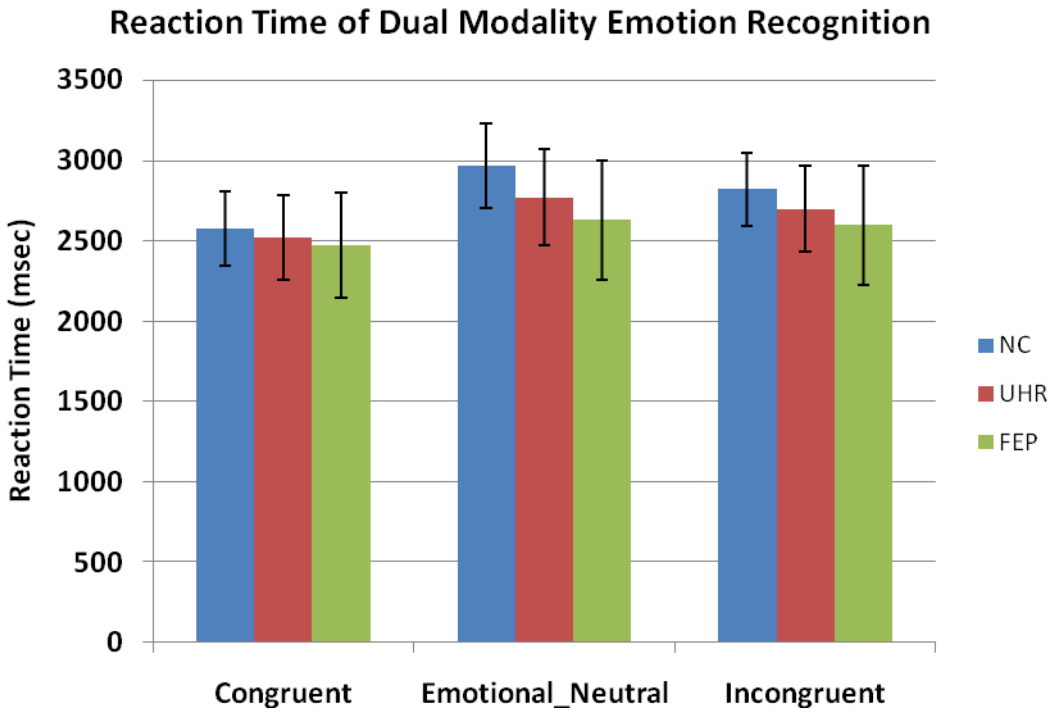


Figure 7.2 (b)



*Emotional-Neutral category shows average performance across Emotional (voice) – Neutral (face) bimodal stimuli.

Figure 7.2 Graph showing (a) mean accuracy, and (b) mean reaction time for each group with respect to emotional congruency.

7.3.3 Functional MRI

7.3.3.1 Whole brain analyses

Average effect of task

Whole brain analysis shows an average effect of task across groups in the bilateral superior and middle temporal gyrus, bilateral inferior frontal gyrus, bilateral fusiform gyrus, bilateral middle occipital gyrus, bilateral angular gyrus, left precentral and post-central gyrus, posterior cingulate gyrus and bilateral cerebellum (FWE $p<0.05$; see Figure 7.3).

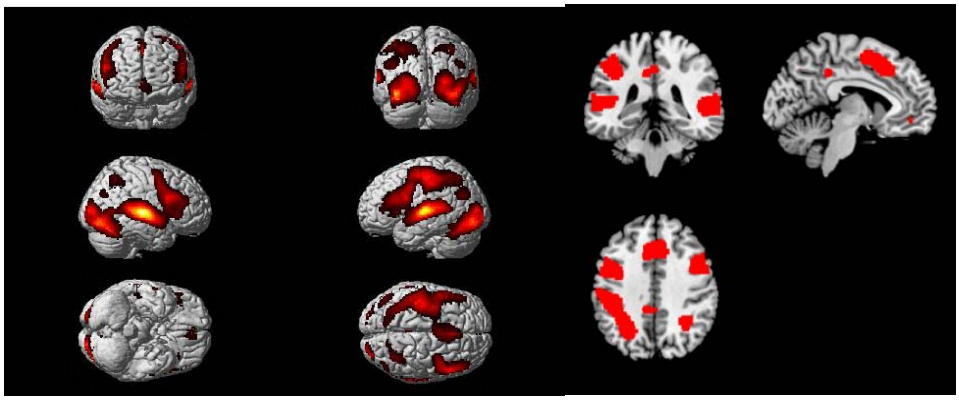


Figure 7.3 SPM showing peak BOLD activation for average effect of task across groups.

Main effect of congruency

There was increased activation during incongruent trials relative to congruent trials in the right supplementary motor area and the bilateral dorsal lateral prefrontal cortex (FWE, $p<0.05$, cluster corrected; see Figure 7.4).

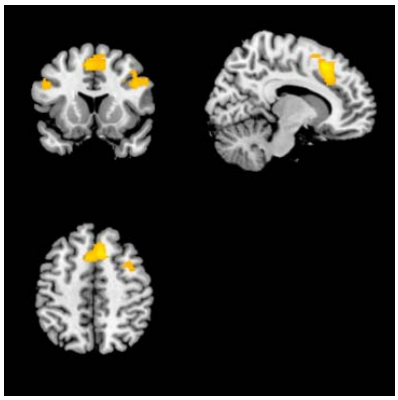


Figure 7.4 SPM showing peak BOLD activation for incongruent more than congruent conditions across groups (FWE, p<0.05, cluster corrected).

7.3.3.2 EFFECT OF GROUP

FEP>NC

There was increased activation in FEP relative to NC in the right ppSTS extending to the right angular gyrus (p<0.05, FWE cluster size corrected, see Figure 7.5(a)).

FEP>UHR

There was increased activation in FEP relative to UHR in the superior medial frontal gyrus (p<0.05, FWE, cluster size corrected, see Figure 7.5(b)).

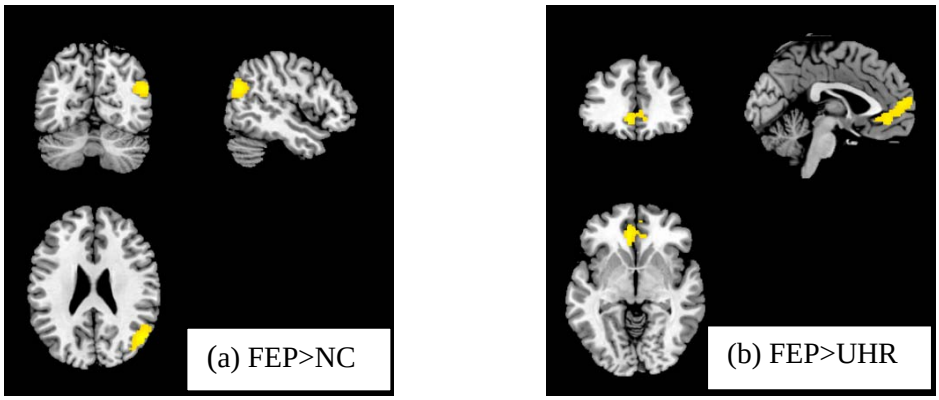


Figure 7.5 SPM showing peak BOLD activation for group effect across congruency, (a) FEP>NC (b) FEP>UHR.

.

	No. of voxels	x	y	z	*Z*	p value Cluster corrected
NC<FEP						
Right angular gyrus/ppSTS	472	50	-66	26	4.28	0.014
UHR<FEP						
Superior medial frontal	570	-6	40	-6	4.06	0.006
		-2	58	10	3.67	

Table 7.1 Coordinates showing group effects (p<0.05, FWE cluster corrected)

7.3.3.3 GROUP × TASK INTERACTION

A group-by-congruency interaction was seen in the left precuneus/PCC (FWE, Cluster corrected, p<0.05, see Figure 7.6(a) and Figure 7.6(b)). In this region FEP patients showed higher activation than both HC and UHR during incongruent trials but not during congruent trials.

A group-by-congruency interaction was also seen in the middle/posterior cingulate cortex and extended to post-central gyrus (Cluster corrected, $p<0.05$, see Figure 7.6(a)). In these regions FEP patients showed higher activation than UHR during incongruent relative to congruent trials (See Figure 7.6(c)).

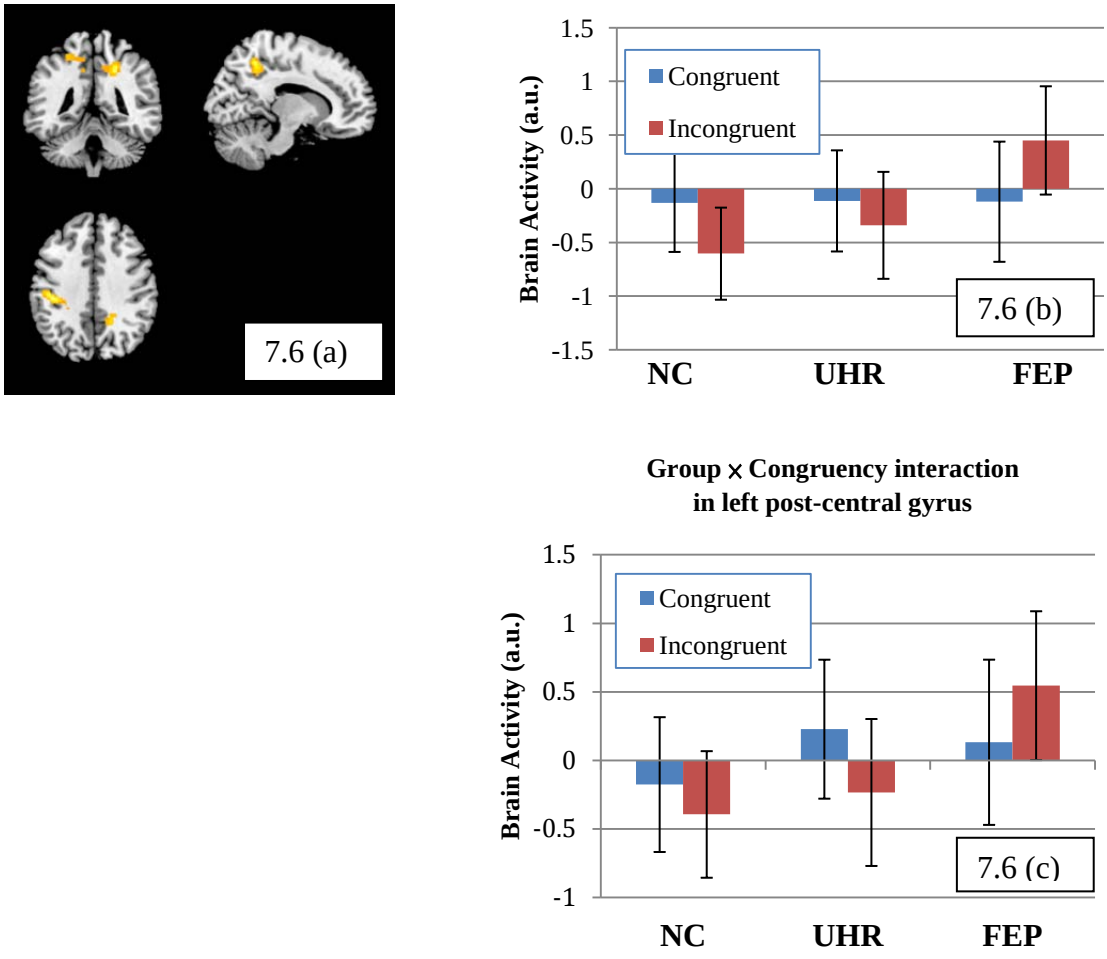


Figure 7.6 Graph (a) and Plots showing peak BOLD activation level of interactions between group level and congruency: (b) right precuneus (c) left post-central gyrus.

Activated brain regions	No. of voxels	x	y	z	*Z* value	*p* value (cluster-corrected)
Group × Task interaction						
Right Precuneus	516	8	-52	50	4.38	0.01
		20	-50	42	4.29	
Left Precuneus		-2	-52	52	3.65	
Middle cingulate gyrus extended to post central gyrus	329	-40	-28	36	4.26	0.046
		-24	-46	56	3.68	
		-10	-40	42	3.60	

Table 7.2 Coordinates showing group × task interactions (p<0.05, FWE cluster corrected)

7.3.3.4 REGION OF INTEREST ANALYSIS

Group Effect

FEP patients had significantly higher activation in the right ppSTS than HC (t=4.02, *p*=0.02) across all trials, mainly driven by less deactivation in FEP patients during incongruent trials (see Figure 7.7).

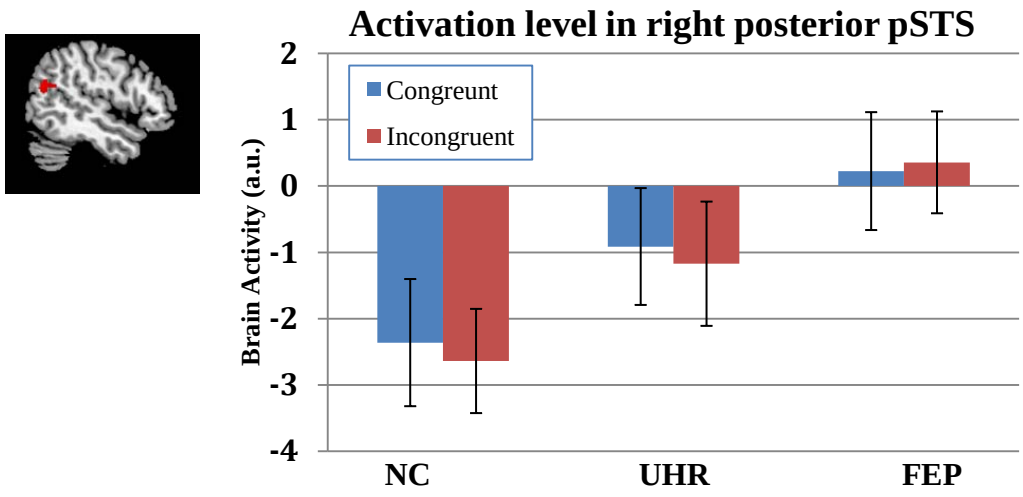


Figure 7.7 Graph showing group effect in right posterior pSTS

Group × Task Interaction

An interaction between HC and FEP patients was seen in both the apSTS (t=4.37, *p*=0.005, right; *t*=3.79, *p*=0.03, left) and right lentiform gyrus/amgydala area (*t*=3.87, *p*=0.03). In the apSTS, the pattern of lower activation for incongruent trials relative to congruent trials in HC was reversed in FEP patients (see Figure 7.8). In the right lentiform gyrus/amgydala significantly increased activation for incongruent trials in FEP patients was not seen in NC (see Figure 7.9). All other interactions were non-significant. All the results reported in this section were uncorrected, p<0.005.

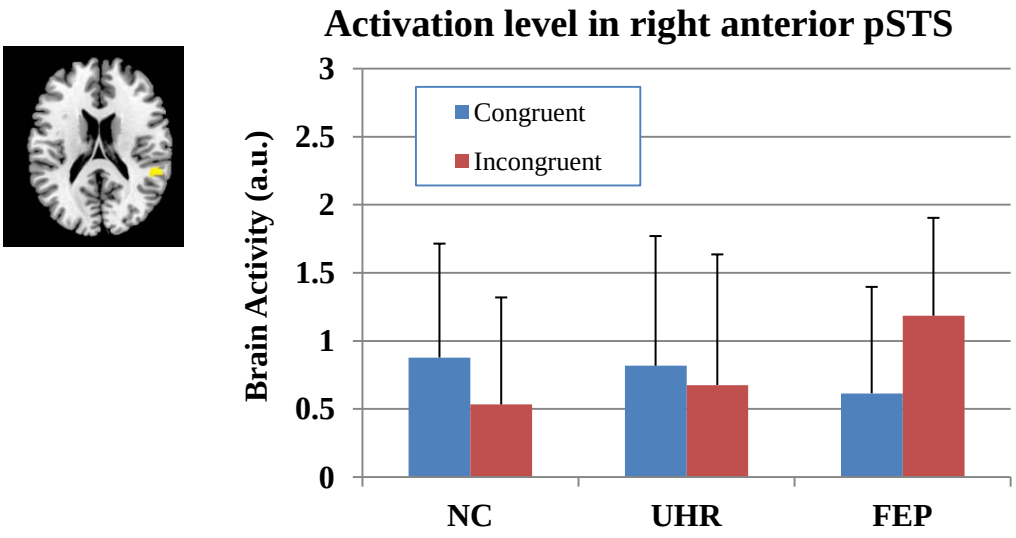


Figure 7.8 Graph showing group effect in posterior pSTS

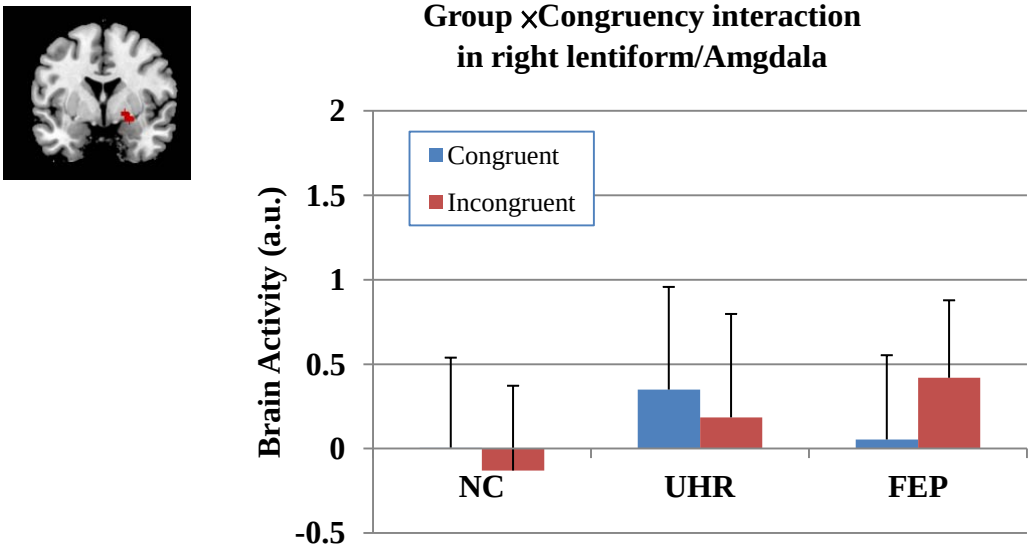


Figure 7.9 Graph showing group effect in Right lentiform neucleus/Amygdala

7.3.4 Correlation Analysis

Using the peak values of parameters extracted from the region of interests that showed significant effects between the groups (i.e. right ppSTS, bilateral apSTS and amygdala), correlation analyses with behavioural performance and clinical symptoms were performed.

7.3.4.1 Task Performance

Across all subjects, during incongruent trials, lower activation in the right ppSTS was associated with longer reaction times (r=-0.29, p=0.03).

7.3.4.2 Clinical symptoms

Higher activation in right ppSTS was associated with higher levels of negative symptoms across clinical groups (r=0.38, p=0.031) during incongruent trials.

7.4 Discussion

The current study aimed to investigate the neural substrate of emotional MSI and if activation in regions responsible for emotional MSI was altered in UHR and FEP groups relative to healthy controls.

7.4.1 Behavioural Results

In line with hypothesis 7.1(a), FEP patients were less accurate during emotional congruent trials, in particular fear trials, which is similar to their performance in the single modality task. As predicted, the performance of UHR subjects was intermediate between those of HC and the FEP patients (although not significantly different from either group). These results suggest that emotional recognition is impaired in first episode rather than prodromal psychosis, although subtle alterations may also be observed in UHR subjects. Furthermore, the pattern of impaired performance was similar across sensory modality, and regardless of the congruency of the trials. However, contrary to hypothesis 7.1(b), as the group by congruency interaction was non-significant, the facilitating effect for congruent relative to incongruent emotional stimuli was similar across the three groups. Overall, the behavioural results suggest that FEP patients had difficulties processing multisensory emotional information, but do not support a specific integration deficit for audiovisual emotional information.

7.4.2 Whole brain analysis

Task-related brain activation (across congruent and incongruent trials) was seen in facial (bilateral occipital facial area, bilateral fusiform gyrus) and prosodic processing areas (bilateral superior temporal gyrus), areas responsible for integration (bilateral posterior STG/STS), and emotion-related neural networks (medial frontal area, anterior and posterior cingulate gyrus, bilateral inferior frontal gyrus). Overall the task activated areas associated with audiovisual integration, social cognition, executive/attentional networks and emotional processing. The subregions of STS (bilateral apSTS) responsible for audiovisual integration (Kreifelts et al., 2009) were activated to a similar extent during both congruent and incongruent trials in all three groups, suggesting the automatic nature of the process. A main effect of congruency across groups was seen in left supplementary motor area, and to a lesser extent, bilateral dorsolateral prefrontal cortices. The increased activation in these areas during incongruent trials was probably due to the demands in cognitive control and attention of the task (Ridderinkhof et al., 2004, Nachev et al., 2008), in order to initiate the button press action or, alternatively, inhibit the response plan (and not to press the button).

However, unlike the HC and UHR subjects, FEP patients failed to deactivate the precuneus/PCC areas during incongruent trials relative to congruent trials. The precuneus is part of the default mode network, and this finding is in line with a number of previous studies showing that patients with psychosis are less able to deactivate the default network when the task load increases (Salgado-Pineda et al., 2011, Schneider et al., 2011, Guerrero-Pedraza et al., 2012). This may reflect the attentional impairments in

patients with psychosis as it has been shown that deactivation in the precuneus is associated with increased concentration in healthy subjects (Weissman et al., 2006).

7.4.3 ROI Results

Results from ROI analyses show that neural substrate related to processing visual emotional information (posterior pSTS, angular gyrus) (Kreifelts et al., 2009) was deactivated in HC and UHR subjects during both congruent and incongruent trials. This may reflect the instruction that was given to all participants to base their response on voice prosodies when they detected conflict between voice and face modalities during incongruent trials. The deactivation in posterior pSTS in HC group suggests that they may have allocated less attention (or tried to completely ignore) to the visual modality and instead only attended to prosodic voice information when making their judgements. A similar pattern of deactivation was seen in the UHR group. FEP patients however did not show deactivation in the posterior pSTS during the task. Consequently, this may have impaired the FEP patients' ability to filter out distracting visual information resulting in lower behavioural accuracy in both congruent and non-congruent conditions.

Broadly consistent with hypothesis 7.3, the brain activation in areas associated with audiovisual integration of emotional stimuli (i.e. bilateral apSTS/right amygdala/lentiform nucleus)(Ethofer et al., 2006b, Kreifelts et al., 2009, Witteman et al., 2012) showed task-specific differences during congruent relative to incongruent trials. In both HC and UHR groups, both areas showed deactivation during incongruent

relative to congruent trials, but FEP patients showed a reverse pattern with greater activation during incongruent trials.

These result suggests that FEP patients may be less able to suppress irrelevant information (i.e. non-target visual information) leading to increased activation in the amygdala due to conflicting emotional information.

This could explain why activation in areas associated with audiovisual integration (anterior pSTS) and discrimination (amygdala) of emotional stimuli was relatively increased in FEP patients during incongruent trials, despite impaired behavioural performances. However, since interaction effects did not survive after correction for multiple comparisons, this interpretation of the results is largely speculative.

During incongruent trials, higher activation in the visual sensitive area (ppSTS) was associated with higher negative symptoms across UHR and FEP subjects. This suggests that FEP and UHR subjects who experienced higher levels of negative symptoms were those less able to deactivate the ppSTS, which might be associated with a gaiting deficit to filter out less important or irrelevant information (Louchart-de la Chapelle et al., 2005).

7.4.4 General discussions

Overall, contrary to the study hypotheses, both the behavioural performances and brain activation patterns in the audiovisual integration area (apSTS) during the MERIT did not suggest specific integration deficits in UHR or FEP subjects. However, the pattern of results does suggest an impaired ability to suppress or filter irrelevant/distracting information in FEP patients. A possible explanation for such results is that the audiovisual MSI is automatic regardless of the congruency of the information, but recognition of emotion from the integrated information can be heavily influenced by cognition and attention. The MERIT used in this study not only elicited the neural networks associated with MSI, but also attention and cognitive control networks, especially during incongruent trials. Thus, the accuracy during the MERIT may reflect the ability to selectively focus on the crucial information and inhibiting unnecessary or distracting information. The possibility of a MSI deficit during multisensory emotional processing in FEP still exists. However, the current task was not able to delineate the process from attention and cognitive control deficits in the FEP patients (Ridderinkhof et al., 2004, Louchart-de la Chapelle et al., 2005, Nachev et al., 2008).

Although the result did not support a specific multisensory integration deficit, FEP patients exhibited multisensory processing difficulties which may be associated with a sensory gaiting deficit (Freedman et al., 1987). FEP patients may be less able filter (or ignore) non-target visual information during incongruent trials. This inability to filter out irrelevant information may explain the significantly greater activation in the visual sensitive part of superior temporal area (ppSTS) during multisensory processing in FEP subjects, which may lead to less effective integration of multisensory information and reduced accuracy. The positive association with negative symptoms suggests that a filtering problem may underlie FEP patients' difficulties in processing external

emotional information from multiple sensory modalities, and may contribute to their social deficits and perplexity.

7.5 Limitations

The main limitation for the MERIT was the absence of non-emotional dual-modality trials as a control task. A potential solution for future studies is using the combination of scrambled faces and rearranged/reversed voice clips as a control task. Secondly, the current study directly compared congruent and incongruent trials to demonstrate integration effects. However, the explicit design for the incongruent task requires selective attention. Given the robust selective attention deficit reported in patients with psychosis, this may confound interpretation regarding MSI performances. For future studies, an implicit design or a simpler explicit design (for example, judging the congruency rather than emotion) may better demonstrate the integration effect.

7.6 Conclusion

The findings suggest that FEP patients were impaired relative to HC in multisensory emotion recognition, whereas the UHR group were not significantly less accurate than HC subjects, although numerically less. Inconsistent with the study hypothesis, both the FEP and the UHR groups did not show specific integration deficits. Both groups showed similar degrees of accuracy facilitation effect during congruent relative to incongruent trials, as that seen in HC subjects. The activation patterns acquired by fMRI suggest that the overall decrease in task performance seen in FEP patients may be due to

a failure to deactivate the posterior pSTS (visual) area, which may be associated with less capability to filter out distracting visual information. Moreover generally, a failure to deactivate the precuneus (part of the DMN) may have impaired task performance when the attentional demands were increased during incongruent trials. In contrast, in UHR subjects, the associated neural networks during MSI seemed relatively unaffected, together with similar recognition accuracy, suggesting a relatively intact MSI ability in the UHR phase.

Chapter 8

General Discussion

Patients in the early phase of psychosis show neuroanatomical and neurofunctional abnormalities similar to those in patients with established schizophrenia (Gur et al., 2002a, Hempel et al., 2003, Habel et al., 2004, Williams et al., 2004), including regions that are normally involved in emotional processing (Fusar-Poli et al., 2007b, Smieskova et al., 2010). To date however, few studies have specifically examined the neural substrate of emotional processing in the early phase of psychosis (Seiferth et al., 2008, Gee et al., 2012). The effective integration of emotional information is extremely important for interpersonal interactions and daily social functioning; but the disturbances of integration of multisensory emotional information and the associated neural processes in patients with the early phase of psychosis remain unclear (See Chapters 1, 2). Moreover, there are no studies that have examined the integration of emotional information in people at risk of psychosis.

A number of studies have shown that multisensory integration (MSI) of non-emotional visual and auditory information is disturbed in patients with schizophrenia (For example, de Gelder et al., 2005, Ross et al., 2007), but the cross-modal integration of emotional information has not been widely studied (For example, de Gelder et al., 2003, de Jong et al., 2009) (See Chapter 2, Table 2.1). Furthermore, functional neuroimaging studies are necessary to elucidate the neural substrate underlying multisensory emotional

processing, and to determine if alterations in this substrate are present before the onset of psychosis.

In this context, assessing emotional recognition and integration in the early phase of psychosis, at both behavioural and neurofunctional levels, is important for establishing whether neurofunctional deficits occur in individual sensory modalities and involve MSI, an important cognitive function for complex social activities. The main aim of this thesis was to investigate emotional processing at the unisensory and multisensory level in both UHR subjects and FEP patients using both visual and auditory emotional stimuli. To this end, both emotional and non-emotional studies of MSI in schizophrenia were reviewed to establish the state of the art in established psychosis (Chapter 2). Next, a multisensory emotional recognition and integration task (MERIT) was developed (Chapter 3) and validated in healthy participants (Chapter 4). Behavioural performance and brain activation patterns associated with the unisensory emotional processing were then examined in HC, UHR subjects and FEP patients (Chapters 5 & 6). Finally, task performance and brain activation during the multisensory conditions of the MERIT, comprising congruent and incongruent audiovisual emotional stimuli, were compared between these subject groups (Chapter 7).

8.1 Summary of Task Development

As outlined in chapter 3, a novel multisensory emotional recognition and integration task (MERIT), was developed to examine audiovisual integration of emotional stimuli in healthy volunteers, UHR and FEP subjects. The task adapted dynamic faces with

emotional prosodic voices from established unisensory emotional recognition tasks (DEER-T and DANVA-2, respectively). A pilot study examining the validity of these prosodic clips in a small UK sample is reported in Chapter 3. Significant facilitation and interference effects were demonstrated in healthy volunteers using this new task (see Chapter 4), which are broadly consistent with previous studies presenting simultaneous multisensory sensory stimuli (de Gelder et al., 1999, de Gelder and Vroomen, 2000, Dolan et al., 2001, Collignon et al., 2008, Jeong et al., 2011). However, unlike in previous studies (Wittfoth et al., 2010, Muller et al., 2011), the overall interference effect on reaction times (RT) was non-significant. The RT interference effect was only significant when happy stimuli were the targets, suggesting an emotion-specific effect on RT for emotional MSI (Chapter 4). In sum, these findings confirmed that MERIT engages emotional audiovisual integrative process in healthy subjects, as judged by the manifested facilitation and interference effects on response accuracy. These results are compatible with the presumption that emotional MSI is normally automatic in nature (Muller et al., 2011).

As outlined in chapter 4, correlation analyses were performed between SPQ scores (i.e. psychometric schizotypy) and facilitation and interference effects during the MERIT. Schizotypy is a personality trait associated with cognitive deficits and emotional disturbances similar to those seen in schizophrenia, albeit to a lesser degree. Correlations between SPQ scores and MERIT performance might indicate possible alterations of emotional MSI in people at high risk for psychosis. Indeed, a moderate positive association was observed between SPQ scores and slower RT during multisensory trials but not during unisensory trials. However, this finding might be

attributed to a more general cognitive deficit under a high processing load associated with schizotypy (Giakoumaki, 2012), rather than impaired MSI, as less effective utilisation of multisensory emotional information was also found during incongruent conditions.

8.2 Summary of Main Findings

8.2.1 Are there emotional processing and integration deficits in FEP patients and UHR subjects?

I hypothesised that, relative to healthy controls, UHR and FEP participants would demonstrate impaired recognition of unisensory emotional information presented in visual (facial expression) and auditory (prosodic cues) modalities, as well as during multisensory emotional trials (*hypothesis 1*). In addition, I predicted that emotionally congruent stimuli would increase the accuracy of emotional recognition and reducing reaction times in healthy volunteers, but not (or to a lesser extent) in UHR subjects and FEP patients during the multisensory emotional task, relative to emotionally incongruent multisensory stimuli (*hypothesis 3*).

Collectively, the findings reported in this thesis are in line with *hypothesis 1* and broadly consistent with the notion that emotional processing is altered in the early stage of psychosis (Thompson et al., 2012). As reported in Chapters 5, 6 and 7, FEP patients were less accurate than the HC group in recognising emotional facial expressions

(Chapter 5), prosodic voices (Chapter 6), *and* multisensory emotional stimuli, particularly during negative emotional trials. These findings are consistent with conclusion drawn from previous meta-analyses of emotion recognition in schizophrenia (Hoekert et al., 2007, Kohler et al., 2010) and literature in FEP patients (Thompson et al., 2012), supporting that the emotion processing deficits already exist in the early stage of psychosis. The emotion-specific pattern of the recognition deficits, comprising worse negative emotion recognition relative to positive emotion recognition, and worst during fear identification, was also similar to that seen in the established psychosis (Edwards et al., 2002). Furthermore, since impairments were demonstrated in both unisensory and multisensory trials, the impairments might be related to a common emotion processing deficit across sensory modalities, rather than deficits specific to any sensory modality.

Contrary to *hypothesis 3*, the facilitating effect for congruent relative to incongruent emotional stimuli was similar across HC and FEP groups; the results hence did not support a specific integration deficit for audiovisual emotional information in FEP patients, but rather a general emotion recognition deficit across multisensory trials. These results are inconsistent with those from previous studies in schizophrenia showing an attenuated facilitation effect (de Jong et al., 2009, de Jong et al., 2010), suggesting that the prominent multisensory emotional processing deficits observed in FEP subjects were, at least in our paradigm, not due to a deficit in MSI.

Consistent with *hypothesis 1,* the UHR subjects' task performance was intermediate between those of HC and FEP subjects. During prosodic voice trials, UHR subjects were significantly less accurate than HC, particularly during fear emotion trials. There were no significant performance differences between UHR and FEP subjects during the MERIT. These results partly replicated the findings of Amminger et al. (2012) suggesting the presence of attenuated emotion recognition difficulties in the UHR group, particularly for fear stimuli.

Inconsistent with *hypothesis 3,* facilitation effects associated with multisensory congruent (relative to incongruent trials) were similar in UHR subjects and HC. The existence of an independent integration deficit in UHR subjects is hence not supported; this finding may have been due to methodological limitations, including relatively small sample size, and the task designs that directly compared congruent and incongruent trials to demonstrate integration effects. The results thus might be confounded by deficits in selective attention which required during incongruent trials (see Section 7.5, page 170 and Section 8.5.2, page 192).

8.2.2 Are there alterations in the neural correlates of emotional processing in UHR subjects and FEP patients?

My hypothesis was that, relative to HC, UHR and FEP participants would demonstrate altered activation during unisensory emotional recognition tasks, especially in areas

responsible for decoding facial (fusiform gyrus) and prosodic information (trunk section of superior STG), emotion identification and interpretation (amygdala, inferior frontal gyrus and hippocampal areas) (hypothesis 2). I also hypothesised that UHR and FEP participants would show altered activation in areas responsible for multisensory integration (superior temporal sulcus/gyrus), and emotion identification and interpretation (amygdala, inferior frontal gyrus and hippocampal areas) during the multisensory emotional task (*hypothesis 4*).

In accordance with hypothesis 2 and hypothesis 4, during both unisensory and multisensory tasks, there was altered activation in areas associated with emotional processing and integration in FEP patients (Chapters 5, 6, 7). There were few significant differences between controls and UHR subjects (Chapters 6, 7). UHR subjects mainly showed hypoactivation in the right fusiform gyrus during the unisensory face task (Chapter 5). UHR subjects showed intermediate alteration of activation between controls and FEP that did not significantly different from both groups in areas responsible for emotional processing and multisensory integration, including fusiform gyrus (Chapter 6), superior temporal gyrus (Chapters 6, 7), and orbital frontal gyrus (Chapters 6).

These findings can be further interpreted according to functions subserved by four areas: Information decoding area, emotional labelling and processing area, areas sensitive to emotional valence, and integration areas.

8.2.2.1 Information decoding areas: Fusiform gyrus and Superior temporal lobe

In line with Seiferth et al. (2008), FEP patients showed reduced activation in the right fusiform gyrus during both facial and prosodic emotion discrimination task relative to HCs, concomitant with lower task accuracy (Chapter 5). Reduced activation in the fusiform gyrus during emotional processing in FEP patients might be associated with impaired decoding of facial information to retrieve emotional content/information. Reduced activation in the fusiform gyrus during the prosodic task further suggested hypo-activation across sensory modality in this region during emotional processing in FEP patients. This is compatible with numerous reports of structural and functional abnormalities in the fusiform gyrus in schizophrenia (e.g., Quintana et al., 2003, Marwick and Hall, 2008) and in first-episode schizophrenia (e.g. Lee et al., 2002).

Patients with schizophrenia who have positive symptoms, in particular auditory hallucinations, have difficulties identifying emotional prosodies (Alba-Ferrara et al., 2012) and tend to misattribute their own distorted speech to an external source (Johns et al., 2001, Johns et al., 2006). These impairments are also present in the UHR population (Johns et al., 2010, Amminger et al., 2012a, Amminger et al., 2012b). Superior temporal gyrus (STG), in particular the posterior region, is responsible for decoding prosodic information. In healthy participants, correct identification of self-generated speech is also associated with activation in the STG (Fu et al., 2006, Allen et al., 2007). Dysfunctional posterior STG activation has also been associated with defective processing of emotional prosody in patients with schizophrenia who experience hallucinations (Alba-Ferrara et al., 2012). Our finding that FEP participants were less accurate and had hypoactivation in STG during the voice prosody task (Chapter 6)

suggest that dysfunction of superior temporal area, especially the posterior part, may underlie the presence of speech processing abnormalities that could contribute to the development of hallucinations.

In UHR subjects, reduced activation in right fusiform gyrus was similar to that seen in FEP patients during the facial emotion discrimination task. Together with the numerically decreased accuracy, reduced activation in this region may be associated with very early decoding difficulties in people at risk of psychosis. Despite the absence of significant activation differences in the middle superior temporal gyrus, an area responsible for prosodic decoding, a significant decrease in accuracy during the prosodic voice task suggests that prosodic decoding problems may not be the primary cause of the deficit in prosodic voice recognition in UHR subjects. The neurofunctional alteration in the later recognition and interpretation stage may provide further explanation for the reduced accuracy (will be discussed in 8.2.2.3, page 181).

8.2.2.2 Emotional identification and discrimination areas:

During the prosodic task, decreased activation in FEP patients was observed in the left amygdala and other emotional processing areas including left insular area, left thalamus, and medial orbital frontal lobe (Chapter 6). A trend for reduced activation in the bilateral inferior frontal and right insular areas was also seen. These results suggest an overall hypoactivation of emotional identification and discrimination related neural network, in particular the left hemisphere. Such finding is in line with literature that consistently reports dysfunction of amygdala and related emotional circuits in patients with schizophrenia (Lee et al., 2002, Li et al., 2010, Tian et al., 2011).

Unexpectedly, no differences in amygdala activation were observed between FEP patients and NC during dynamic facial stimuli (Chapter 5). Given the decreased recognition accuracy in FEP patients relative to controls during dynamic face trials and reduced activation in facial decoding area (i.e., the fusiform gyrus). The results suggest that areas associated with decoding facial expressions, rather than emotional identification and discrimination, might have accounted for such recognition deficits in the early stage of psychosis.

To date, structural or functional alterations of amygdala function have not been reported in UHR subjects. I did not find differences in activation between UHR subjects and controls during both unisensory and multisensory emotional tasks. In line with the findings of Rasetti et al. (2009), in UHR subjects, activation in left amygdala was significantly higher than that in FEP patients but not significantly altered relative to HCs during the prosodic task. This may be the result of higher levels of inter-subject variability due to clinical heterogeneity in the UHR group and relatively small sample size. An alternative explanation is that rather than being an endophenotypic trait, hypoactivation of the amygdala is related to the disease state and only observed in those experiencing full-blown psychotic episodes (Rasetti et al., 2009).

8.2.2.3 Areas sensitive to emotional valence

Valence-specific activation in the caudate has been reported in healthy volunteers (Carretie et al., 2009). In particular, the caudate body is known to be sensitive to

negative emotional stimuli. This effect was verified in the healthy controls in the present thesis (see Section 6.3.2.3, page 134). Contrarily, during both facial and prosodic tasks, there were no activation differences between NC and FEP patients in the striatum or caudate area. Nevertheless, a group by task interaction in the caudate body during the prosodic task revealed a pattern of greater activation for positive relative to negative emotions in UHR subjects, a reversal of the activation pattern in healthy controls (Carretie et al., 2009). UHR subjects also had significantly reduced fearful prosody recognition. This may be due to a positive-negative emotion misattribution brought about by aberrant caudate body activation in UHR group and implicate an interpretation error in UHR subjects, in contrast with a global emotional processing problem in FEP group.

8.2.2.4 Integration associated areas: middle and posterior STS

The sub-regions of STS (bilateral bifurcation areas of pSTS, or in the current thesis, anterior pSTS) responsible for audiovisual integration (Kreifelts et al., 2009) were activated to a similar extent during both congruent and incongruent trials in NC, UHR and FEP groups. This result is compatible with the supposition that the MSI process is automatic in nature and was similar in all three groups. Despite similar activation in all three groups, FEP patients showed greater activation during incongruent relative to congruent trials, a reversal of the pattern in HC group. This suggests that increased effort may be required to integrate conflicting audiovisual information in FEP patients and might be associated with a filtering impairment (See Chapter 7) that will be further described below.

The neural substrate related to processing visual emotional information (posterior terminal ascending branch of pSTS, angular gyrus) (Kreifelts et al., 2009) in STS did not show significant deactivation in FEP patients during the dual modality task as in HC and UHR subjects. Together with lower recognition accuracy across congruency conditions, this finding suggests an inability to filter out distracting information (all subjects were instructed to make judgments according to the emotion in the voice). Since in FEP patients an absence of deactivation was also observed in the PCC, a more generalised deficit brought about by a failure to deactivate DMN network, rather than specific emotional processing deficits, could have accounted for the finding.

The pattern of activation/deactivation in integration-related areas in the UHR group during both congruent and incongruent trials was similar to that seen in the HC group. The results suggest that although numerically less accurate than the NC group, the ability to selectively allocate attention when making judgements remained intact in UHR subjects. Together with the findings from single modality prosodic trials, these findings suggest that in UHR subjects, brain activation during decoding and integration was relatively intact, but that valence misattribution might underlie reduced accuracy, particularly for negative emotional stimuli.

8.3 Emotional Processing and Integration difficulties in psychosis: A Speculative model

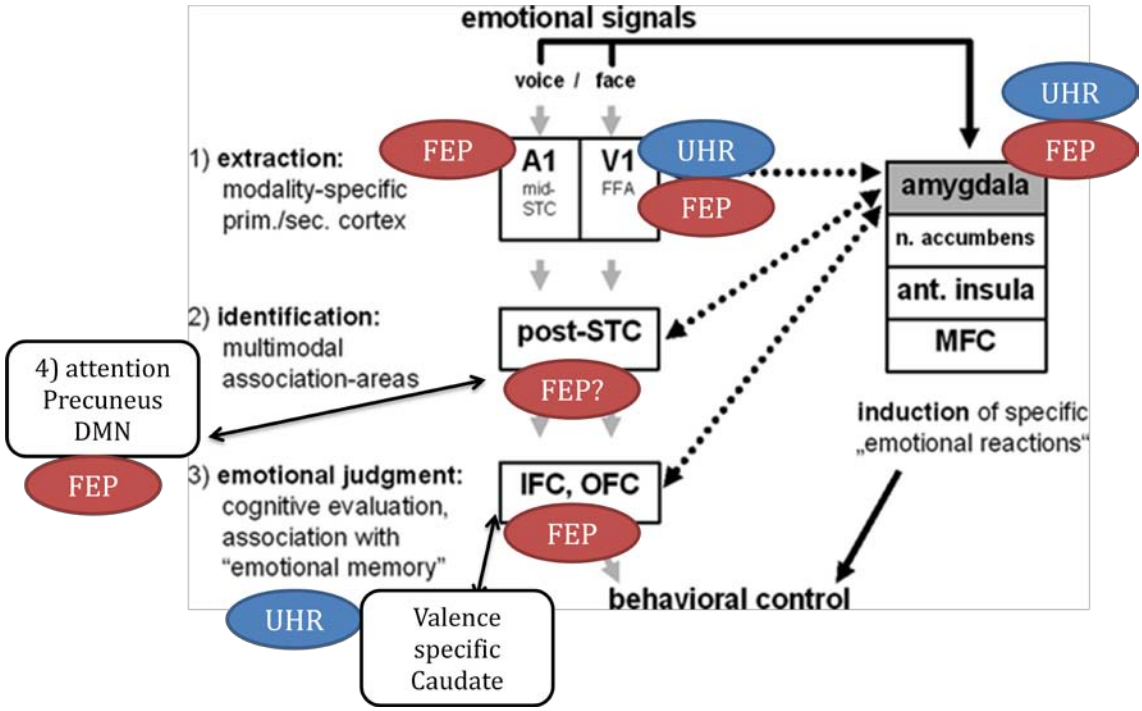


Figure 8.1 A speculative model for emotional processing deficit in UHR and FEP (Adapted from Wildgruber et al., 2009)

The integration of information is an automatic process, which occurs early in the stream of cognitive processing. However, it is possible that during multisensory emotional processing, recognition, discrimination and labeling of emotional valence do not occur simultaneously during integration. Instead, similar to emotional processing in unisensory emotional recognition, a coarse labeling of the stimuli as 'emotional' might also occur during sensory integration, just after the unisensory decoding stage of facial and prosodic information. The posterior STG/STS, known to be crucial for integration of information from visual and auditory modalities, may also be the area responsible for

initial labeling of the integrated percept as emotional before an efferent signal is sent to frontal and limbic regions where evaluation and categorization occur.

This speculative model thus suggests that the posterior STS/STG plays an important role in multisensory emotional processing in several ways: i) to integrate visual and auditory information to form a unified percept; ii) to coarsely label the percept as emotional (or non-emotional), and iii) to modulate the weighting of auditory and visual information and filter out the unnecessary information. The decoded auditory and visual information in fusiform gyrus and middle STS/STG is then sent to the posterior STS/STG, where the information is integrated and forms the coarse emotional percept. The primordial percept is then sent forward to the anterior brain and limbic system for detailed processing and interpretation. Modality-specific attentional modulation can occur in the areas sensitive to different modalities along STS. According to the current thesis, the visually sensitive area (i.e., posterior pSTS) may change the weightings of visual and auditory information, and consequently influence the content of the integrated percept to be processed and the accuracy of recognition.

Deficits in multisensory integration have been demonstrated in patients with schizophrenia (See reviews in Chapter 2). However, a specific deficit in 'emotional integration' is difficult to establish because previous studies did not clearly delineate different stages of cognitive processing of information with emotional content. Behaviourally, most existing paradigms have only examined performance accuracy, reaction time, or the 'efficiency' (the accuracy/RT trade-off), disregarding the

distinction between the three functional levels in multisensory processing of emotional information: early sensory processing; integration and emotional percept formation; recognition, discrimination and interpretation. Impaired recognition performance could have arisen from deficits from any or all of the three processing levels. Another important theoretical construct to be considered is, as shown in Chapter 7, the impaired ability to discount or filter out unnecessary information which might interfere with the processing stream, particularly during the integration and recognition/interpretation stages when input from multiple brain area need to be processed at the same time.

My results demonstrate that FEP patients have reduced activation during emotional trials relative to neutral or near-neutral trials in both face (Chapter 5) and prosodic processing areas (Chapter 6), suggesting an early sensory processing deficits likely at the decoding phase. During the dual modality tasks, although recognition accuracy differed between NC and FEP patients, levels of activation did not differ between the two groups in the integration area of pSTS (near bifurcation of the posterior branches, see Chapter 7). This suggests that activation in this area during multisensory trials could not have accounted for the low accuracy in the FEP and UHR group.

It should be emphasized that activation in the posterior STS during multisensory trials may not necessarily result in successful integration of sensory information into an emotional percept. First, a trend of less activation of the integration area in posterior STS in FEP subjects during congruent multisensory stimuli can still be seen, suggesting insufficient integrating activation in FEP patients. Second, a comparable degree of

activation for multisensory stimuli in the FEP and UHR groups might not indicate a similar degree of effective integration. The integrated information being signaled forward to the emotional recognition areas in FEP may be less effectively filtered. More importantly, conflicting or ambiguous information could be wrongly labeled as 'emotional', resulting in "messy" integrated percept hard to recognise and interpret.

The deactivation seen in healthy participants in the face sensitive area of the STS was possibly due to selective sensory filtering under current task instruction as a strategy to reduce the ambiguity inherent in incongruent multisensory trials. Nevertheless, emotionally conflicting information may still evoke automatic integration mechanisms, only to result in a unified percept which is judged to be less emotional (Muller et al., 2011). Following the instructions given before the task, healthy participants may have adopted a strategy whereby they ignored information from auditory (prosodic) stimuli, thereby suppressing brain activation in the STS sub-region sensitive to emotional visual information. The ability to effectively suppress activation in this region may have allowed participants to focus on auditory stimuli only, resulting in less ambiguous percept formation. This initial filtering and integration process is likely to be more cognitively demanding and involves less emotional interpretation. Consequently, it is less likely that an emotional reaction within the individual will be triggered (Figure 8.3, 'Induction'). As a result, the areas responsible for triggering an emotional response (such as amygdala) were not strongly activated. On the contrary, in the FEP patients, the potentially unresolved percept (due to insufficient sensory filtering) being signaled to emotional processing and recognition regions was still coarsely labeled. This in turn may have resulted in greater activation in the amygdala and associated brain areas

during recognition and interpretation. The failure to deactivate region in the DMN may also have contributed to filtering deficits in FEP patients, resulting in insufficient suppression of unwanted and unrelated brain processes.

Overall, a picture emerges that suggests that FEP patients might have processing difficulties at all three stages emotional MSI; not only the early processing stage (Chapter 5), but also the integration/attention modulation stage (Chapter 7), and emotional recognition and interpretation stage (Chapters 5, 6, 7). In UHR subjects, some early processing problem may exist, especially for configurations in facial expressions (Chapter 5 and Chapter 6). However, the main impairment may be at the recognition and interpretation stage, where they may misattribute positive and negative emotional stimuli. For a summary of the findings, see Figure 8.10.

8.4 Study Features

8.4.1 Development and validation of a novel multisensory tasks

Facial emotion recognition is the most frequently used paradigm to examine both emotional and social cognition deficits in clinical groups. However, the performance in any single modality is insufficient to generalise to the difficulties they may have in those complex real life situation when dealing with information from multiple sensory modalities. By employing a novel tool, examining emotional processing ability across

visual and auditory sensory modality, the results of the current thesis contribute to the understanding of emotional processing difficulties in the psychosis spectrum, and provide some insight into how these difficulties might develop and manifest in individuals at high risk of or already in the early stage of psychosis.

8.4.2 The joint examination of both UHR subjects and FEP patients

Investigating emotional processing in people thought to be at risk of imminent psychosis, and those who recently experienced a frank psychotic episode in the same study has a number of advantages. Firstly, from a general perspective, many of the confounding factors associated with investigations of established psychosis are minimised, including; effects related to the duration of illness, and effects of long-term antipsychotic medication. Secondly, studying those in the early phase of psychosis may contribute to our understanding of neurobiological mechanisms that lead to the onset of illness. This could inform early intervention strategies, especially in relation to emotional and social difficulties.

Thirdly, the study format allowed direct comparisons between FEP patients, UHR and HC subjects. For example, whilst UHR subjects differed significantly from HCs in terms of recognition accuracy for visual, auditory and audiovisual stimuli, no corresponding alteration was evident with respect to neurofunction. This finding indicates that, even though the presentation of social and emotional processing

difficulties in the UHR group look similar to those seen in FEP patients, the underlying brain mechanisms may differ. Alternatively, the inherent heterogeneity within the UHR subjects may have 'diluted' any effects that might of be observable in a true prodromal group.

8.4.3 Examining both emotion recognition and audiovisual integration

The majority of studies conducted to date have been unisensory in their approach and typically contrasted either schizophrenia/FEP and HC subjects. By applying this relatively novel approach using multiple sensory modalities, the results generated provide a more holistic account of emotional processing. Further investigation of the integration process of audiovisual information is not only of mechanistic importance, but is more ecologically close to the social and cognitive abilities used in everyday life.

8.5 Study Limitations

Overall, there are a number of limitations related to the investigation that need to be discussed. These are summarised as follows; i) sample size and distribution, ii) task design, iii) cross-sectional versus longitudinal; iv) exposure to antipsychotic medication.

8.5.1 Sample size and distribution

The main limitation of the present study is the relatively small sample size. In Chapter 3, though the sample size (n=25, HC) is large enough to demonstrate facilitation and interference effects, it was not optimal to test for associations between psychometric schizotypy and multisensory recognition and integration. Furthermore, a distribution skewed to the lower range of the SPQ scores suggested that schizoptypal traits were low in this population, making interpretation of the results difficult.

In terms of neurofunctional analysis, a recent report suggested optimal group sizes, for the purposes of classical inference, of 16-35 (Friston, 2012). After the removal of subjects with extreme movement artefacts that were identified during preprocessing, the sample size of each group in the current thesis is around the lower boundary of this recommended group size (i.e. 20, 15, 17 for HC, UHR, FEP, respectively). These sample sizes were sufficient to demonstrate neurofunctional effects in FEP patients, but may not have been optimal to fully demonstrate more subtle changes in the UHR group, particularly given their clinical heterogeneity that are related to several potential outcomes rather than being unique to psychotic disorders (Fusar-Poli et al., 2013a).

To retrospectively examine the feasibility of the task, the effect size for amygdala activation elicited by emotional faces (Chapter 5) and voices task (Chapter 6) in healthy participants (n=20) was calculated. Overall the effect sizes are moderate (0.57 and 0.52 in face paradigm and voice paradigm, respectively), suggesting that the task would be feasible for demonstrating amygdala responses in healthy participants with a reasonable sample size (around 25). However, though the sample size of 25 in each group has been

estimated to demonstrate overall group differences (See Chapter 5), it may still not be enough to demonstrate modest between-group activation differences in amygdala.

Since these comparisons were performed using relatively modest sample sizes, it is not possible to draw a definite conclusion regarding the potential differences between groups, particularly with respect to the negative findings UHR subjects. Due to the time constraints of a PhD project and the unpredictable speed of recruitment, obtaining optimal sample sizes in each group for statistical comparisons were not feasible.

The fact that the majority of FEP subjects were older than HC and UHR subjects could also be another limitation of the study. Whilst the age difference was small and might not have an impact on accuracy and/or brain activation, it remains a potential confound with respect to the findings. In all the analyses age has been statistically controlled, but at the same time the statistics were less robust to demonstrate potential differences between groups.

8.5.2 Task design

There are several potential limitations related to the design of the MERIT. Firstly, dynamic facial emotion recognition is a relatively cognitively-demanding task to perform under a time constraint. As such, for those subjects with a higher degree of cognitive and functional decline, their performance could have been impaired as a result of a general cognitive impairment, particularly attentional deficits, rather than a specific effect of impaired emotional processing. Although the three groups did not show

significant difference in estimated FIQ, but the FEP patients were numerically lower than the other two groups and their attentional function has not been independently assessed. Secondly, the prosodic voices task did not include an emotionally neutral voice condition, because there were no neutral prosodic stimuli in the original design of DANVA2 (See Section 6.5, page 142). In my analyses I took low-intensity stimuli as the baseline condition to contrast against high- intensity voices (See Section 6.2.3.2, page 125), which may not have been optimal. Thirdly, directly comparing congruent and incongruent trials to demonstrate integration effects may confound interpretations regarding MSI with attentional function (see Section 7.5, page 170). These limitations make it difficult to exclude the effect of a general cognitive impairment on task performance.

Nevertheless, the regions activated by the task in HC were similar to those described in the literature for emotional processing and emotional integration (Witteman et al., 2012). Furthermore, social cognition is a complex process involving both emotional and cognitive process. As such it may not be possible to entirely separate cognitive and emotional processes during social cognition tasks.

8.5.3 Cross sectional design versus longitudinal perspective

The cross sectional design of the current study was not able to provide information about changes over time, or provide predictive information about transition to psychosis

in UHR subjects. A longitudinal perspective is necessary to delineate changes over time and establish how these relate to clinical outcomes. However, years or more in order to define clinical outcome (Fusar-Poli et al., 2012), and large baseline samples so that there are sufficient numbers of subjects in the different outcome subgroups for statistical comparisons. Conducting such a study was not feasible as a PhD project.

8.5.4 Antipsychotics use in FEP

The majority of FEP patients recruited to the current study were not antipsychotic naïve. To date, the precise effects of antipsychotic medication on neurofunction remain unclear (Fusar-Poli et al., 2007a, Navari and Dazzan, 2009). Nevertheless, a difference in medication exposure is a potential confounder when interpreting behavioural or activation differences between FEP and HC, or FEP and UHR. Although most of the studies in schizophrenia suggest that the dosage of antipsychotics was not a major confounding influence on performance during emotion recognition, it remains possible that medication may have still have contributed to the performance of multisensory integration. Furthermore, the influence of antipsychotic medication on neural activity during emotional processing remains unclear.

8.6 Final Conclusion and Future Work

To conclude, the results presented in this thesis are in part consistent with the notion that, relative to HCs, FEP patients show neurofunctional alterations in emotional processing regions that are qualitatively similar to those previously observed in

schizophrenia patients. However, in UHR subjects, such alterations may be less pronounced than in FEP patients.

Though the FEP and UHR group did not show a significant alteration in the brain areas associated with integration (i.e. anterior pSTS), the results demonstrates that FEP patients failed to deactivate areas that may have helped to filter irrelevant visual stimuli, and areas associated with the DMN which is likely to have impaired their ability to perform the task at a cognitive level. A speculative model proposes that pSTS/STG is important for integrating and forming a unified emotional percept, and that activation in this region can be modulate by modality-specific attention. The failure of this modulation may partially explain the difficulties identifying multisensory emotional stimuli in FEP patients.

The importance and potential contribution of the thesis are: 1) the development of emotional behavioural and fMRI paradigms involving both visual and auditory modalities which have greater ecological validity. 2) Demonstrating unisensory and multisensory emotional processing difficulties in the early and prodromal stage of psychosis, at both behavioural and neurofunctional levels. 3) The findings provide a potential link between clinical symptoms, behavioural performances and brain activation in the early and prodromal stage of psychosis. 4) The provision of a speculative neurofunctional model for the emotional processing deficits in the early and prodromal stage of psychosis.

The positive association between brain activation and clinical symptoms sheds light on possible mechanism that may contribute to the manifestation of symptoms, and also raises the possibility of establishing neurofunctional biomarkers that could be used to identify UHR subjects who a have higher risk of frank psychosis, a prospect which could be investigated in future prospective and longitudinal studies.

APPENDIX

A: Publications Arising

“Multisensory emotion integration of dynamic faces and prosodic voices: individual differences in healthy volunteers with varying level of schizotypy”, Huai-Hsuan Tseng, Jonathan P Roiser, Sunjeev Kamboj, Bradley Platt, Jeffery Dalton, Philip McGuire, & Paul Allen. (Submitted to Cognition and Emotion)

“A systemic review of non-emotional and emotional multisensory integration deficit in Schizophrenia”, Huai-Hsuan Tseng, Matthijs Bossong, Gemma Modinos, Kuan-Ming Chen, Philip McGuire, & Paul Allen. (In Preparation)

REFERENCES

Abbott GR, Green MJ (2012) Facial affect recognition and schizotypal personality characteristics. Early Interv Psychiatry.

Addington J, Addington D (1998) Facial affect recognition and information processing in schizophrenia and bipolar disorder. Schizophr Res 32:171-181.

Addington J, Addington D (2008) Symptom remission in first episode patients. Schizophr Res 106:281-285.

Addington J, Piskulic D, Perkins D, Woods SW, Liu L, Penn DL (2012) Affect recognition in people at clinical high risk of psychosis. Schizophr Res 140:87-92.

Adolphs R (2002) Neural systems for recognizing emotion. Curr Opin Neurobiol 12:169-177.

Adolphs R (2008) Fear, faces, and the human amygdala. Curr Opin Neurobiol 18:166-172.

Adolphs R, Tranel D, Damasio H, Damasio AR (1995) Fear and the human amygdala. J Neurosci 15:5879-5891.

Alba-Ferrara L, Fernyhough C, Weis S, Mitchell RLC, Hausmann M (2012) Contributions of emotional prosody comprehension deficits to the formation of auditory verbal hallucinations in schizophrenia. Clinical Psychology Review 32:244-250.

Allen P, Amaro E, Fu CH, Williams SC, Brammer MJ, Johns LC, McGuire PK (2007) Neural correlates of the misattribution of speech in schizophrenia. Br J Psychiatry 190:162-169.

Amminger GP, Schafer MR, Klier CM, Schlogelhofer M, Mossaheb N, Thompson A, Bechdolf A, Allott K, McGorry PD, Nelson B (2012a) Facial and vocal affect perception in people at ultra-high risk of psychosis, first-episode schizophrenia and healthy controls. Early Interv Psychiatry 6:450-454.

Amminger GP, Schafer MR, Papageorgiou K, Klier CM, Schlogelhofer M, Mossaheb N, Werneck-Rohrer S, Nelson B, McGorry PD (2012b) Emotion recognition in individuals at clinical high-risk for schizophrenia. Schizophr Bull 38:1030-1039.

Arsalidou M, Morris D, Taylor MJ (2011) Converging evidence for the advantage of dynamic facial expressions. Brain Topogr 24:149-163.

Bach DR, Buxtorf K, Grandjean D, Strik WK (2009) The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. Psychol Med 39:927-938.

Baum KM, Nowicki S (1998) Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. J Nonverbal Behav 22:89-107.

Bediou B, Franck N, Saoud M, Baudouin JY, Tiberghien G, Dalery J, d'Amato T (2005) Effects of emotion and identity on facial affect processing in schizophrenia. Psychiatry Res 133:149-157.

Bertelson P, Radeau M (1981) Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. Percept Psychophys 29:578-584.

Blake R, Shiffrar M (2007) Perception of human motion. Annu Rev Psychol 58:47-73.

Blanchard JJ, Collins LM, Aghevli M, Leung WW, Cohen AS (2011) Social anhedonia and schizotypy in a community sample: the Maryland longitudinal study of schizotypy. Schizophr Bull 37:587-602.

Bleuler E (1911) Dementia Praecox or the Group of Schizophrenias.

Borgwardt SJ, McGuire PK, Aston J, Berger G, Dazzan P, Gschwandtner U, Pfluger M, D'Souza M, Radue EW, Riecher-Rossler A (2007) Structural brain abnormalities in individuals with an at-risk mental state who later develop psychosis. Br J Psychiatry Suppl 51:s69-75.

Brittain P, Ffytche DH, McKendrick A, Surguladze S (2010) Visual processing, social cognition and functional outcome in schizophrenia. Psychiatry Res 178:270-275.

Broome MR, Matthiasson P, Fusar-Poli P, Woolley JB, Johns LC, Tabraham P, Bramon E, Valmaggia L, Williams SC, Brammer MJ, Chitnis X, McGuire PK (2009) Neural correlates of executive function and working memory in the 'at-risk mental state'. Br J Psychiatry 194:25-33.

Broome MR, Woolley JB, Johns LC, Valmaggia LR, Tabraham P, Gafoor R, Bramon E, McGuire PK (2005) Outreach and support in south London (OASIS): implementation of a clinical service for prodromal psychosis and the at risk mental state. Eur Psychiatry 20:372-378.

Brown LA, Cohen AS (2010) Facial emotion recognition in schizotypy: the role of accuracy and social cognitive bias. J Int Neuropsychol Soc 16:474-483.

Bruck C, Kreifelts B, Kaza E, Lotze M, Wildgruber D (2011) Impact of personality on the cerebral processing of emotional prosody. Neuroimage 58:259-268.

Calvert GA, Hansen PC, Iversen SD, Brammer MJ (2001) Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. Neuroimage 14:427-438.

Cannon TD, Cadenhead K, Cornblatt B, Woods SW, Addington J, Walker E, Seidman LJ, Perkins D, Tsuang M, McGlashan T, Heinssen R (2008) Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry 65:28-37.

Carretie L, Rios M, de la Gandara BS, Tapia M, Albert J, Lopez-Martin S, Alvarez-Linera J (2009) The striatum beyond reward: caudate responds intensely to unpleasant pictures. Neuroscience 164:1615-1622.

Chan RC, Li H, Cheung EF, Gong QY (2010) Impaired facial emotion perception in schizophrenia: a meta-analysis. Psychiatry Res 178:381-390.

Chapman LJ, Chapman JP, Kwapil TR, Eckblad M, Zinser MC (1994) Putatively psychosis-prone subjects 10 years later. J Abnorm Psychol 103:171-183.

Choe CS, Welch RB, Gilford RM, Juola JF (1975) Ventriloquist Effect - Visual Dominance or Response Bias. Perception & Psychophysics 18:55-60.

Claridge G, Robinson DL, Birchall P (1985) Psychophysiological Evidence of Psychoticism in Schizophrenics Relatives. Pers Indiv Differ 6:1-10.

Colavita FB (1974) Human Sensory Dominance. Perception & Psychophysics 16:409-412.

Collignon O, Girard S, Gosselin F, Roy S, Saint-Amour D, Lassonde M, Lepore F (2008) Audio-visual integration of emotion expression. Brain Res 1242:126-135.

Compton MT, Bercu Z, Bollini A, Walker EF (2006) Factor structure of the Neurological Evaluation Scale in a predominantly African American sample of patients with schizophrenia, unaffected relatives, and non-psychiatric controls. Schizophrenia Research 84:365-377.

de Boer-Schellekens L, Stekelenburg JJ, Maes JP, Van Gool AR, Vroomen J (2013) Sound improves diminished visual temporal sensitivity in schizophrenia. Acta Psychol (Amst).

de Gelder B, Bocker KBE, Tuomainen J, Hensen M, Vroomen J (1999) The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. Neuroscience Letters 260:133-136.

de Gelder B, Vroomen J (2000) Bimodal emotion perception: integration across separate modalities, cross-modal perceptual grouping or perception of multimodal events? Cognition Emotion 14:321-324.

de Gelder B, Vroomen J, Annen L, Masthof E, Hodiamont P (2003) Audio-visual integration in schizophrenia. Schizophr Res 59:211-218.

de Gelder B, Vroomen J, de Jong SJ, Masthoff ED, Trompenaars FJ, Hodiamont P (2005) Multisensory integration of emotional faces and voices in schizophrenics. Schizophr Res 72:195-203.

de Jong JJ, Hodiamont PP, de Gelder B (2010) Modality-specific attention and multisensory integration of emotions in schizophrenia: reduced regulatory effects. Schizophr Res 122:136-143.

de Jong JJ, Hodiamont PP, Van den Stock J, de Gelder B (2009) Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. Schizophr Res 107:286-293.

Demenescu LR, Kortekaas R, den Boer JA, Aleman A (2010) Impaired attribution of emotion to facial expressions in anxiety and major depression. PLoS One 5:e15058.

Demjaha A, Valmaggia L, Stahl D, Byrne M, McGuire P (2010) Disorganization/Cognitive and Negative Symptom Dimensions in the At-Risk Mental State Predict Subsequent Transition to Psychosis. Schizophr Bull.

Dolan RJ, Morris JS, de Gelder B (2001) Crossmodal binding of fear in voice and face. Proc Natl Acad Sci U S A 98:10006-10010.

Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. Neuron 57:11-23.

Eack SM, Mermon DE, Montrose DM, Miewald J, Gur RE, Gur RC, Sweeney JA, Keshavan MS (2010) Social cognition deficits among individuals at familial high risk for schizophrenia. Schizophr Bull 36:1081-1088.

Edwards J, Jackson HJ, Pattison PE (2002) Emotion recognition via facial expression and affective prosody in schizophrenia: a methodological review. Clin Psychol Rev 22:789-832.

Edwards J, Pattison PE, Jackson HJ, Wales RJ (2001) Facial affect and affective prosody recognition in first-episode schizophrenia. Schizophrenia Research 48:235-253.

Egerton A, Borgwardt SJ, Tognin S, Howes OD, McGuire P, Allen P (2011) An overview of functional, structural and neurochemical imaging studies in

individuals with a clinical high risk for psychosis. Neuropsychiatry-Lond 1:477-493.

Ethofer T, Anders S, Erb M, Droll C, Royen L, Saur R, Reiterer S, Grodd W, Wildgruber D (2006a) Impact of voice on emotional judgment of faces: an event-related fMRI study. Hum Brain Mapp 27:707-714.

Ethofer T, Pourtois G, Wildgruber D (2006b) Investigating audiovisual integration of emotional signals in the human brain. In: Progress in Brain Research, vol. Volume 156 (S. Anders, G. E. M. J. J. K. and Wildgruber, D., eds), pp 345-361: Elsevier.

Ferri F, Costantini M, Salone A, Di Iorio G, Martinotti G, Chiarelli A, Merla A, Di Giannantonio M, Gallese V (2013) Upcoming tactile events and body ownership in schizophrenia. Schizophr Res.

Foa EB, Kozak MJ (1986) Emotional processing of fear: exposure to corrective information. Psychol Bull 99:20-35.

Foucher JR, Lacambre M, Pham BT, Giersch A, Elliott MA (2007) Low time resolution in schizophrenia Lengthened windows of simultaneity for visual, auditory and bimodal stimuli. Schizophr Res 97:118-127.

Freedman R, Adler LE, Gerhardt GA, Waldo M, Baker N, Rose GM, Drebing C, Nagamoto H, Bickford-Wimer P, Franks R (1987) Neurobiological studies of sensory gating in schizophrenia. Schizophr Bull 13:669-678.

Freeman D, Garety PA (2003) Connecting neurosis and psychosis: the direct influence of emotion on delusions and hallucinations. Behav Res Ther 41:923-947.

Friston K (2012) Ten ironic rules for non-statistical reviewers. NeuroImage 61:1300-1310.

Friston K, Ashburner J, Frith C (1995) Spatial registration and normalization of images. In: Human Brain Mapping.

Fu CH, Vythelingum GN, Brammer MJ, Williams SC, Amaro E, Jr., Andrew CM, Yaguez L, van Haren NE, Matsumoto K, McGuire PK (2006) An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. Cereb Cortex 16:969-977.

Fusar-Poli P, Bonoldi I, Yung AR, Borgwardt S, Kempton MJ, Valmaggia L, Barale F, Caverzasi E, McGuire P (2012) Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch Gen Psychiatry 69:220-229.

Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rossler A, Schultze-Lutter F, Keshavan M, Wood S, Ruhrmann S, Seidman LJ, Valmaggia L, Cannon T, Velthorst E, De Haan L, Cornblatt B, Bonoldi I, Birchwood M, McGlashan T, Carpenter W, McGorry P, Klosterkotter J, McGuire P, Yung A (2013a) The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA psychiatry 70:107-120.

Fusar-Poli P, Broome MR, Matthiasson P, Williams SCR, Brammer M, McGuire PK (2007a) Effects of acute antipsychotic treatment on brain activation in first episode psychosis: an fMRI study. Eur Neuropsychopharmacol 17:492-500.

Fusar-Poli P, Broome MR, Matthiasson P, Woolley JB, Mechelli A, Johns LC, Tabraham P, Bramon E, Valmaggia L, Williams SC, McGuire P (2011) Prefrontal function at presentation directly related to clinical outcome in people at ultrahigh risk of psychosis. Schizophr Bull 37:189-198.

Fusar-Poli P, Byrne M, Badger S, Valmaggia LR, McGuire PK (2013b) Outreach and support in south London (OASIS), 2001-2011: ten years of early diagnosis and treatment for young individuals at high clinical risk for psychosis. Eur Psychiatry 28:315-326.

Fusar-Poli P, Perez J, Broome M, Borgwardt S, Placentino A, Caverzasi E, Cortesi M, Veggiotti P, Politi P, Barale F, McGuire P (2007b) Neurofunctional correlates of vulnerability to psychosis: a systematic review and meta-analysis. Neurosci Biobehav Rev 31:465-484.

Fusar-Poli P, Placentino A, Carletti F, Landi P, Allen P, Surguladze S, Benedetti F, Abbamonte M, Gasparotti R, Barale F, Perez J, McGuire P, Politi P (2009) Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. J Psychiatry Neurosci 34:418-432.

Gee DG, Karlsgodt KH, van Erp TG, Bearden CE, Lieberman MD, Belger A, Perkins DO, Olvet DM, Cornblatt BA, Constable T, Woods SW, Addington J, Cadenhead KS, McGlashan TH, Seidman LJ, Tsuang MT, Walker EF, Cannon TD, Consortium N (2012) Altered age-related trajectories of amygdala-prefrontal circuitry in adolescents at clinical high risk for psychosis: a preliminary study. Schizophr Res 134:1-9.

Grice GR, Spiker VA (1979) Speed-Accuracy Tradeoff in Choice Reaction-Time - within Conditions, between Conditions, and between Subjects. Perception & Psychophysics 26:118-126.

Guerrero-Pedraza A, McKenna PJ, Gomar JJ, Sarro S, Salvador R, Amann B, Carrion MI, Landin-Romero R, Blanch J, Pomarol-Clotet E (2012) First-episode psychosis is characterized by failure of deactivation but not by hypo- or hyperfrontality. Psychol Med 42:73-84.

Gur RE, McGrath C, Chan RM, Schroeder L, Turner T, Turetsky BI, Kohler C, Alsop D, Maldjian J, Ragland JD, Gur RC (2002a) An fMRI study of facial emotion processing in patients with schizophrenia. Am J Psychiatry 159:1992-1999.

Gur RE, McGrath C, Chan RM, Schroeder L, Turner T, Turetsky BI, Kohler C, Alsop D, Maldjian J, Ragland JD, Gur RC (2002b) An fMRI study of facial emotion processing in patients with schizophrenia. Am J Psychiat 159:1992-1999.

Haas BW, Omura K, Constable RT, Canli T (2006) Interference produced by emotional conflict associated with anterior cingulate activation. Cogn Affect Behav Neurosci 6:152-156.

Habel U, Klein M, Shah NJ, Toni I, Zilles K, Falkai P, Schneider F (2004) Genetic load on amygdala hypofunction during sadness in nonaffected brothers of schizophrenia patients. Am J Psychiatry 161:1806-1813.

Hempel A, Hempel E, Schonknecht P, Stippich C, Schroder J (2003) Impairment in basal limbic function in schizophrenia during affect recognition. Psychiatry Res 122:115-124.

Hoekert M, Bais L, Kahn RS, Aleman A (2008) Time course of the involvement of the right anterior superior temporal gyrus and the right fronto-parietal operculum in emotional prosody perception. PLoS One 3:e2244.

Hoekert M, Kahn RS, Pijnenborg M, Aleman A (2007) Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. Schizophr Res 96:135-145.

Hooker C, Park S (2002) Emotion processing and its relationship to social functioning in schizophrenia patients. Psychiat Res 112:41-50.

Horan WP, Blanchard JJ (2003) Neurocognitive, social, and emotional dysfunction in deficit syndrome schizophrenia. Schizophr Res 65:125-137.

Horan WP, Blanchard JJ, Clark LA, Green MF (2008) Affective traits in schizophrenia and schizotypy. Schizophr Bull 34:856-874.

Howes OD, Bose SK, Turkheimer F, Valli I, Egerton A, Valmaggia LR, Murray RM, McGuire P (2011) Dopamine synthesis capacity before onset of psychosis: a prospective [18F]-DOPA PET imaging study. Am J Psychiatry 168:1311-1317.

Jacklin DL, Goel A, Clementino KJ, Hall AW, Talpos JC, Winters BD (2012) Severe Cross-Modal Object Recognition Deficits in Rats Treated Sub-Chronically with NMDA Receptor Antagonists are Reversed by Systemic Nicotine: Implications for Abnormal Multisensory Integration in Schizophrenia. Neuropsychopharmacology.

Jardri R, Pins D, Bubrovszky M, Lucas B, Lethuc V, Delmaire C, Vantyghem V, Despretz P, Thomas P (2009) Neural functional organization of hallucinations in schizophrenia: multisensory dissolution of pathological emergence in consciousness. Conscious Cogn 18:449-457.

Jeong JW, Diwadkar VA, Chugani CD, Sinsoongsud P, Muzik O, Behen ME, Chugani HT, Chugani DC (2011) Congruence of happy and sad emotion in music and faces modifies cortical audiovisual activation. Neuroimage 54:2973-2982.

Johns LC, Allen P, Valli I, Winton-Brown T, Broome M, Woolley J, Tabraham P, Day F, Howes O, Wykes T, McGuire P (2010) Impaired verbal self-monitoring in individuals at high risk of psychosis. Psychol Med 40:1433-1442.

Johns LC, Gregg L, Allen P, McGuire PK (2006) Impaired verbal self-monitoring in psychosis: effects of state, trait and diagnosis. Psychol Med 36:465-474.

Johns LC, Rossell S, Frith C, Ahmad F, Hemsley D, Kuipers E, McGuire PK (2001) Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. Psychol Med 31:705-715.

Jones PB, Bebbington P, Foerster A, Lewis SW, Murray RM, Russell A, Sham PC, Toone BK, Wilkins S (1993) Premorbid social underachievement in schizophrenia. Results from the Camberwell Collaborative Psychosis Study. Br J Psychiatry 162:65-71.

Kamboj SK, Joye A, Bisby JA, Das RK, Platt B, Curran HV (2012) Processing of facial affect in social drinkers: a dose-response study of alcohol using dynamic emotion expressions. Psychopharmacology (Berl).

Kay S, Opler, L., Fiszbein, A. (1987) The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. In: Schizophr Bull.

Kee KS, Green MF, Mintz J, Brekke JS (2003) Is emotion processing a predictor of functional outcome in schizophrenia? Schizophr Bull 29:487-497.

Kee KS, Horan WP, Mintz J, Green MF (2004) Do the siblings of schizophrenia patients demonstrate affect perception deficits? Schizophr Res 67:87-94.

Kendler KS, McGuire M, Gruenberg AM, O'Hare A, Spellman M, Walsh D (1993) The Roscommon Family Study. III. Schizophrenia-related personality disorders in relatives. Arch Gen Psychiatry 50:781-788.

Kohler CG, Bilker W, Hagendoorn M, Gur RE, Gur RC (2000) Emotion recognition deficit in schizophrenia: association with symptomatology and cognition. Biol Psychiatry 48:127-136.

Kohler CG, Martin EA (2006) Emotional processing in schizophrenia. Cogn Neuropsychiatry 11:250-271.

Kohler CG, Turner TH, Bilker WB, Brensinger CM, Siegel SJ, Kanes SJ, Gur RE, Gur RC (2003) Facial emotion recognition in schizophrenia: intensity effects and error pattern. Am J Psychiatry 160:1768-1774.

Kohler CG, Walker JB, Martin EA, Healey KM, Moberg PJ (2010) Facial emotion perception in schizophrenia: a meta-analytic review. Schizophr Bull 36:1009-1019.

Koppen C, Spence C (2007) Seeing the light: exploring the Colavita visual dominance effect. Exp Brain Res 180:737-754.

Kosaka H, Omori M, Murata T, Iidaka T, Yamada H, Okada T, Takahashi T, Sadato N, Itoh H, Yonekura Y, Wada Y (2002) Differential amygdala response during facial recognition in patients with schizophrenia: an fMRI study. Schizophrenia Research 57:87-95.

Kreifelts B, Ethofer T, Shiozawa T, Grodd W, Wildgruber D (2009) Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. Neuropsychologia 47:3059-3066.

Kucharska-Pietura K, David AS, Masiak M, Phillips ML (2005) Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. Br J Psychiatry 187:523-528.

Laroi F, Fonteneau B, Mourad H, Raballo A (2010) Basic emotion recognition and psychopathology in schizophrenia. J Nerv Ment Dis 198:79-81.

Lee CU, Shenton ME, Salisbury DF, Kasai K, Onitsuka T, Dickey CC, Yurgelun-Todd D, Kikinis R, Jolesz FA, McCarley RW (2002) Fusiform gyrus volume reduction in first-episode schizophrenia: a magnetic resonance imaging study. Arch Gen Psychiatry 59:775-781.

Leentjens AF, Wielaert SM, van Harskamp F, Wilmink FW (1998) Disturbances of affective prosody in patients with schizophrenia; a cross sectional study. Journal of Neurology, Neurosurgery & Psychiatry 64:375-378.

Leitman DI, Foxe JJ, Butler PD, Saperstein A, Revheim N, Javitt DC (2005) Sensory contributions to impaired prosodic processing in schizophrenia. Biol Psychiatry 58:56-61.

Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, Nierenberg J, Jalbrzikowski M, Lim KO, Javitt DC (2007) The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. Am J Psychiatry 164:474-482.

Leppanen JM, Tenhunen M, Hietanen JK (2003) Faster choice-reaction times to positive than to negative facial expressions - The role of cognitive and motor processes. J Psychophysiol 17:113-123.

Li H, Chan RC, McAlonan GM, Gong QY (2010) Facial emotion processing in schizophrenia: a meta-analysis of functional neuroimaging data. Schizophr Bull 36:1029-1039.

Liebowitz MR (1987) Delineating Social Phobia. Brit J Psychiat 150:718-719.

Louchart-de la Chapelle S, Levillain D, Menard JF, Van der Elst A, Allio G, Haouzir S, Dollfus S, Campion D, Thibaut F (2005) P50 inhibitory gating deficit is correlated with the negative symptomatology of schizophrenia. Psychiatry Res 136:27-34.

Lovibond PF, Lovibond SH (1995) The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 33:335-343.

Mandal MK, Pandey R, Prasad AB (1998) Facial expressions of emotions and schizophrenia: A review. Schizophrenia Bull 24:399-412.

Marwick K, Hall J (2008) Social cognition in schizophrenia: a review of face processing. British medical bulletin 88:43-58.

McGurk H, MacDonald J (1976) Hearing lips and seeing voices. Nature 264:746-748.

Mechelli A, Riecher-Rossler A, Meisenzahl EM, Tognin S, Wood SJ, Borgwardt SJ, Koutsouleris N, Yung AR, Stone JM, Phillips LJ, McGorry PD, Valli I, Velakoulis D, Woolley J, Pantelis C, McGuire P (2011) Neuroanatomical abnormalities that predate the onset of psychosis: a multicenter study. Arch Gen Psychiatry 68:489-495.

Meehl PE (1962) Schizotaxia, Schizotypy, Schizophrenia. American Psychologist 17:827-838.

Miller J (1982) Divided attention: Evidence for coactivation with redundant signals. Cognitive Psychology 14:247-279.

Miller J (1986) Timecourse of coactivation in bimodal divided attention. Attention, Perception, & Psychophysics 40:331-343.

Mitchell RL, Crow TJ (2005) Right hemisphere language functions and schizophrenia: the forgotten hemisphere? Brain 128:963-978.

Mitchell RL, Elliott R, Barry M, Cruttenden A, Woodruff PW (2004) Neural response to emotional prosody in schizophrenia and in bipolar affective disorder. Br J Psychiatry 184:223-230.

Modinos G, Ormel J, Aleman A (2010) Altered activation and functional connectivity of neural systems supporting cognitive control of emotion in psychosis proneness. Schizophr Res 118:88-97.

Muller VI, Cieslik EC, Kellermann TS, Eickhoff SB (2013a) Crossmodal emotional integration in major depression. Soc Cogn Affect Neurosci.

Muller VI, Cieslik EC, Laird AR, Fox PT, Eickhoff SB (2013b) Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization. Front Hum Neurosci 7:268.

Muller VI, Habel U, Derntl B, Schneider F, Zilles K, Turetsky BI, Eickhoff SB (2011) Incongruence effects in crossmodal emotional integration. Neuroimage 54:2257-2266.

Muller VI, Kellermann TS, Seligman SC, Turetsky BI, Eickhoff SB (2012) Modulation of Affective Face Processing Deficits in Schizophrenia by Congruent Emotional Sounds. Soc Cogn Affect Neurosci.

Nachev P, Kennard C, Husain M (2008) Functional role of the supplementary and pre-supplementary motor areas. Nat Rev Neurosci 9:856-869.

Najolia GM, Cohen AS, Minor KS (2011) A laboratory study of affectivity in schizotypy: subjective and lexical analysis. Psychiatry Res 189:233-238.

Navari S, Dazzan P (2009) Do antipsychotic drugs affect brain structure? A systematic and critical review of MRI findings. Psychological medicine.

Nowicki S, Duke MP (1994) Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. J Nonverbal Behav 18:9-35.

Ochsner KN, Hughes B, Robertson ER, Cooper JC, Gabrieli JD (2009) Neural systems supporting the control of affective and cognitive conflicts. J Cogn Neurosci 21:1842-1855.

Ogawa S, Lee TM (1990) Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 16:9-18.

Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci U S A 87:9868-9872.

Pan YJ, Chen SH, Chen WJ, Liu SK (2009) Affect recognition as an independent social function determinant in schizophrenia. Compr Psychiatry 50:443-452.

Park JY, Gu BM, Kang DH, Shin YW, Choi CH, Lee JM, Kwon JS (2010) Integration of cross-modal emotional information in the human brain: an fMRI study. Cortex 46:161-169.

Pauly K, Seiferth NY, Kellermann T, Ruhrmann S, Daumann B, Backes V, Klosterkotter J, Shah NJ, Schneider F, Kircher TT, Habel U (2010) The interaction of working memory and emotion in persons clinically at risk for psychosis: An fMRI pilot study. Schizophrenia Research 120:167-176.

Pearl D, Yodashkin-Porat D, Katz N, Valevski A, Aizenberg D, Sigler M, Weizman A, Kikinzon L (2009) Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. Compr Psychiatry 50:186-192.

Penn DL, Combs DR, Ritchie M, Francis J, Cassisi J, Morris S, Townsend M (2000) Emotion recognition in schizophrenia: further investigation of generalized versus specific deficit models. J Abnorm Psychol 109:512-516.

Penn DL, Corrigan PW, Bentall RP, Racenstein JM, Newman L (1997) Social cognition in schizophrenia. Psychological Bulletin 121:114-132.

Phillips LJ, Yung AR, McGorry PD (2000) Identification of young people at risk of psychosis: validation of Personal Assessment and Crisis Evaluation Clinic intake criteria. Aust N Z J Psychiatry 34 Suppl:S164-169.

Phillips LK, Seidman LJ (2008) Emotion processing in persons at risk for schizophrenia. Schizophr Bull 34:888-903.

Pinkham A, Penn D, Wangelin B, Perkins D, Gerig G, Gu H, Lieberman J (2005) Facial emotion perception and fusiform gyrus volume in first episode schizophrenia. Schizophrenia Research 79:341-343.

Pinkham AE, Penn DL, Perkins DO, Lieberman J (2003) Implications for the neural basis of social cognition for the study of schizophrenia. Am J Psychiatry 160:815-824.

Poole JH, Tobias FC, Vinogradov S (2000) The functional relevance of affect recognition errors in schizophrenia. J Int Neuropsychol Soc 6:649-658.

Pourtois G, de Gelder B, Bol A, Crommelinck M (2005) Perception of facial expressions and voices and of their combination in the human brain. Cortex 41:49-59.

Quintana J, Wong T, Ortiz-Portillo E, Marder SR, Mazziotta JC (2003) Right lateral fusiform gyrus dysfunction during facial information processing in schizophrenia. Biol Psychiatry 53:1099-1112.

Raballo A, Nelson B, Thompson A, Yung A (2011) The Comprehensive Assessment of At-Risk Mental States: From mapping the onset to mapping the structure. Schizophrenia Research 127:107-114.

Rachman S (1980) Emotional processing. Behav Res Ther 18:51-60.

Raine A (1991) The SPQ: a scale for the assessment of schizotypal personality based on DSM-III-R criteria. Schizophr Bull 17:555-564.

Rasetti R, Mattay VS, Wiedholz LM, Kolachana BS, Hariri AR, Callicott JH, Meyer-Lindenberg A, Weinberger DR (2009) Evidence that altered amygdala activity in schizophrenia is related to clinical state and not genetic risk. Am J Psychiatry 166:216-225.

Reske M, Habel U, Kellermann T, Backes V, Jon Shah N, von Wilmsdorff M, Gaebel W, Zilles K, Schneider F (2009) Differential brain activation during facial emotion discrimination in first-episode schizophrenia. J Psychiatr Res 43:592-599.

Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS (2004) Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn 56:129-140.

Roiser JP, Howes OD, Chaddock CA, Joyce EM, McGuire P (2012) Neural and Behavioral Correlates of Aberrant Salience in Individuals at Risk for Psychosis. Schizophr Bull.

Ross LA, Saint-Amour D, Leavitt VM, Molholm S, Javitt DC, Foxe JJ (2007) Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. Schizophr Res 97:173-183.

Rossell SL, Boundy CL (2005) Are auditory-verbal hallucinations associated with auditory affective processing deficits? Schizophr Res 78:95-106.

Ruhrmann S, Schultze-Lutter F, Maier W, Klosterkötter J (2005) Pharmacological intervention in the initial prodromal phase of psychosis. Eur Psychiatry 20:1-6.

Salgado-Pineda P, Fakra E, Delaveau P, McKenna PJ, Pomarol-Clotet E, Blin O (2011) Correlated structural and functional brain abnormalities in the default mode network in schizophrenia patients. Schizophr Res 125:101-109.

Sanders RD, Joo YH, Almasy L, Wood J, Keshavan MS, Pogue-Geile MF, Gur RC, Gur RE, Nimgaonkar VL (2006) Are neurologic examination abnormalities heritable? A preliminary study. Schizophr Res 86:172-180.

Sanders RD, Keshavan MS, Forman SD, Pieri JN, McLaughlin N, Allen DN, van Kammen DP, Goldstein G (2000) Factor structure of neurologic examination abnormalities in unmedicated schizophrenia. Psychiatry Res 95:237-243.

Sass K, Heim S, Sachs O, Straube B, Schneider F, Habel U, Kircher T (2013) Neural correlates of semantic associations in patients with schizophrenia. Eur Arch Psychiatry Clin Neurosci.

Schneider F, Gur RC, Gur RE, Shtasel DL (1995) Emotional processing in schizophrenia: neurobehavioral probes in relation to psychopathology. Schizophrenia Research 17:67-75.

Schneider FC, Royer A, Grosselin A, Pellet J, Barral FG, Laurent B, Brouillet D, Lang F (2011) Modulation of the default mode network is task-dependant in chronic schizophrenia patients. Schizophr Res 125:110-117.

Schröger E, Widmann A (1998) Speeded responses to audiovisual signal changes result from bimodal integration. Psychophysiology 35:755-759.

Seiferth NY, Pauly K, Habel U, Kellermann T, Shah NJ, Ruhrmann S, Klosterkotter J, Schneider F, Kircher T (2008) Increased neural response related to neutral faces in individuals at risk for psychosis. Neuroimage 40:289-297.

Seubert J, Loughead J, Kellermann T, Boers F, Brensinger CM, Habel U (2010) Multisensory integration of emotionally valenced olfactory-visual information in patients with schizophrenia and healthy controls. J Psychiatry Neurosci 35:185-194.

Shea TL, Sergejew AA, Burnham D, Jones C, Rossell SL, Copolov DL, Egan GF (2007) Emotional prosodic processing in auditory hallucinations. Schizophr Res 90:214-220.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 59 Suppl 20:22-33;quiz 34-57.

Shin YW, Na MH, Ha TH, Kang DH, Yoo SY, Kwon JS (2008) Dysfunction in configural face processing in patients with schizophrenia. Schizophr Bull 34:538-543.

Silver H, Shlomo N, Turner T, Gur RC (2002) Perception of happy and sad facial expressions in chronic schizophrenia: evidence for two evaluative systems. Schizophrenia Research 55:171-177.

Smieskova R, Fusar-Poli P, Allen P, Bendfeldt K, Stieglitz RD, Drewe J, Radue EW, McGuire PK, Riecher-Rossler A, Borgwardt SJ (2010) Neuroimaging predictors of transition to psychosis--a systematic review and meta-analysis. Neurosci Biobehav Rev 34:1207-1222.

Smiley JF, Konnova K, Bleiwas C (2012) Cortical thickness, neuron density and size in the inferior parietal lobe in schizophrenia. Schizophr Res 136:43-50.

Smiley JF, Rosoklija G, Mancevski B, Mann JJ, Dwork AJ, Javitt DC (2009) Altered volume and hemispheric asymmetry of the superficial cortical layers in the schizophrenia planum temporale. Eur J Neurosci 30:449-463.

Stekelenburg JJ, Maes JP, Van Gool AR, Sitskoorn M, Vroomen J (2013) Deficient multisensory integration in schizophrenia: an event-related potential study. Schizophr Res 147:253-261.

Stephen JM, Coffman BA, Jung RE, Bustillo JR, Aine CJ, Calhoun VD (2013) Using joint ICA to link function and structure using MEG and DTI in schizophrenia. Neuroimage 83C:418-430.

Stone DB, Urrea LJ, Aine CJ, Bustillo JR, Clark VP, Stephen JM (2011) Unisensory processing and multisensory integration in schizophrenia: a high-density electrical mapping study. Neuropsychologia 49:3178-3187.

Straube B, Green A, Sass K, Kircher T (2013) Superior Temporal Sulcus Disconnectivity During Processing of Metaphoric Gestures in Schizophrenia. Schizophr Bull.

Surguladze SA, Calvert GA, Brammer MJ, Campbell R, Bullmore ET, Giampietro V, David AS (2001) Audio-visual speech perception in schizophrenia: an fMRI study. Psychiatry Res 106:1-14.

Suslow T, Roestel C, Droste T, Arolt V (2003) Automatic processing of verbal emotion stimuli in schizophrenia. Psychiatry Res 120:131-144.

Szycik GR, Munte TF, Dillo W, Mohammadi B, Samii A, Emrich HM, Dietrich DE (2009) Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. Schizophr Res 110:111-118.

Szycik GR, Ye Z, Mohammadi B, Dillo W, Te Wildt BT, Samii A, Frieling H, Bleich S, Munte TF (2013) Maladaptive connectivity of Broca's area in schizophrenia during audiovisual speech perception: An fMRI study. Neuroscience.

Thompson A, Papas A, Bartholomeusz C, Allott K, Amminger GP, Nelson B, Wood S, Yung A (2012) Social cognition in clinical "at risk" for psychosis and first episode psychosis populations. Schizophr Res 141:204-209.

Tian L, Meng C, Yan H, Zhao Q, Liu Q, Yan J, Han Y, Yuan H, Wang L, Yue W, Zhang Y, Li X, Zhu C, He Y, Zhang D (2011) Convergent evidence from multimodal imaging reveals amygdala abnormalities in schizophrenic patients and their first-degree relatives. PLoS One 6:e28794.

Tognin S, Riecher-Rossler A, Meisenzahl EM, Wood SJ, Hutton C, Borgwardt SJ, Koutsouleris N, Yung AR, Allen P, Phillips LJ, McGorry PD, Valli I, Velakoulis D, Nelson B, Woolley J, Pantelis C, McGuire P, Mechelli A (2013) Reduced parahippocampal cortical thickness in subjects at ultra-high risk for psychosis. Psychol Med 1-10.

Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C (2009) The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res 168:242-249.

Tseng HH, Chen SH, Liu CM, Howes O, Huang YL, Hsieh MH, Liu CC, Shan JC, Lin YT, Hwu HG (2013) Facial and prosodic emotion recognition deficits associate with specific clusters of psychotic symptoms in schizophrenia. PLoS One 8:e66571.

Tumkaya S, Karadag F, Oguzhanoglu NK (2012) Neurological soft signs in schizophrenia and obsessive compulsive disorder spectrum. Eur Psychiatry 27:192-199.

Van den Stock J, de Jong SJ, Hodiamont PP, de Gelder B (2011) Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. Soc Neurosci 6:537-547.

van Rijn S, Aleman A, van Diessen E, Berckmoes C, Vingerhoets G, Kahn RS (2005) What is said or how it is said makes a difference: role of the right fronto-parietal operculum in emotional prosody as revealed by repetitive TMS. Eur J Neurosci 21:3195-3200.

Velthorst E, Nieman DH, Klaassen RM, Becker HE, Dingemans PM, Linszen DH, De Haan L (2011) Three-year course of clinical symptomatology in young people at ultra high risk for transition to psychosis. Acta Psychiatr Scand 123:36-42.

Walker EF, Grimes KE, Davis DM, Smith AJ (1993) Childhood precursors of schizophrenia: facial expressions of emotion. Am J Psychiatry 150:1654-1660.

Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. Nat Neurosci 9:971-978.

Welch RB, Warren DH (1980) Immediate Perceptual Response to Intersensory Discrepancy. Psychological Bulletin 88:638-667.

Wildgruber D, Ethofer T, Grandjean D, Kreifelts B (2009) A cerebral network model of speech prosody comprehension. International Journal of Speech-Language Pathology 11:277-281.

Williams LE, Light GA, Braff DL, Ramachandran VS (2010) Reduced multisensory integration in patients with schizophrenia on a target detection task. Neuropsychologia 48:3128-3136.

Williams LM, Das P, Harris AW, Liddell BB, Brammer MJ, Olivieri G, Skerrett D, Phillips ML, David AS, Peduto A, Gordon E (2004) Dysregulation of arousal and amygdala-prefrontal systems in paranoid schizophrenia. Am J Psychiatry 161:480-489.

Witteman J, Van Heuven VJ, Schiller NO (2012) Hearing feelings: a quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. Neuropsychologia 50:2752-2763.

Wittfoth M, Schroder C, Schardt DM, Dengler R, Heinze HJ, Kotz SA (2010) On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. Cereb Cortex 20:383-392.

Yung AR, McGorry PD (1996) The initial prodrome in psychosis: descriptive and qualitative aspects. Aust N Z J Psychiatry 30:587-599.

Yung AR, Nelson B, Stanford C, Simmons MB, Cosgrave EM, Killackey E, Phillips LJ, Bechdolf A, Buckby J, McGorry PD (2008a) Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. Schizophr Res 105:10-17.

Yung AR, Nelson B, Stanford C, Simmons MB, Cosgrave EM, Killackey E, Phillips LJ, Bechdolf A, Buckby J, McGorry PD (2008b) Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. Schizophrenia research 105:10-17.

Yung AR, Phillips LJ, McGorry PD, McFarlane CA, Francey S, Harrigan S, Patton GC, Jackson H (1998) Prediction of psychosis. A step towards indicated prevention of schizophrenia. In: The British journal of psychiatry Supplement, vol. 172, pp 14-20.

Yung AR, Phillips LJ, Yuen HP, Francey SM, McFarlane CA, Hallgren M, McGorry PD (2003a) Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophr Res 60:21-32.

Yung AR, Phillips LJ, Yuen HP, Francey SM, McFarlane CA, Hallgren M, McGorry PD (2003b) Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophrenia research 60:21-32.

Yung AR, Yuen HP, McGorry PD, Phillips LJ, Kelly D, Dell'Olio M, Francey SM, Cosgrave EM, Killackey E, Stanford C, Godfrey K, Buckby J (2005) Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry 39:964-971.